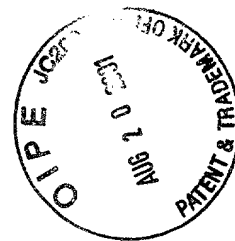


100290 6539260

+



#4

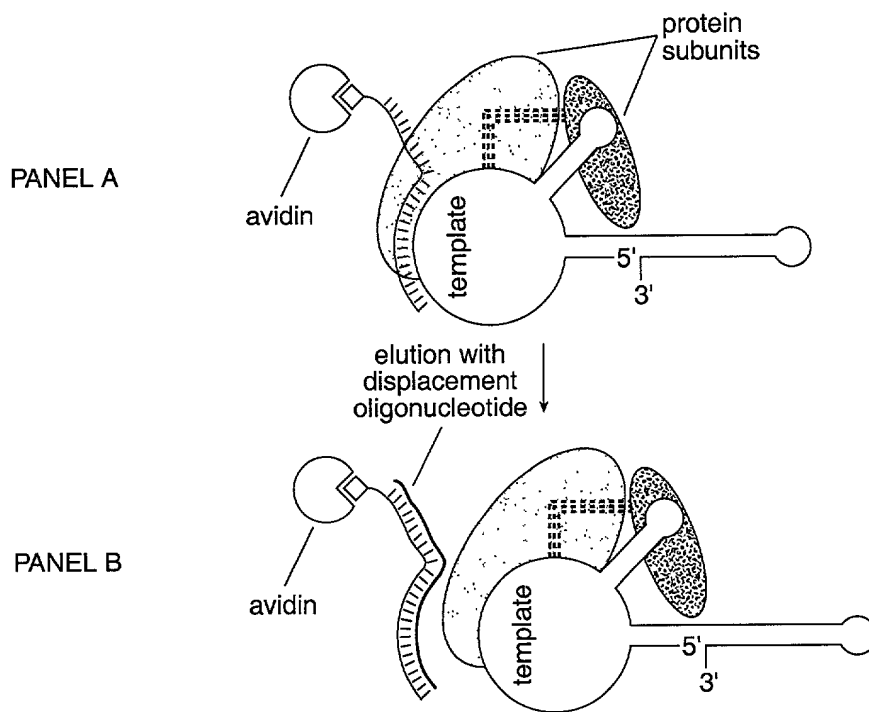


FIG. 1

+

+

2/59

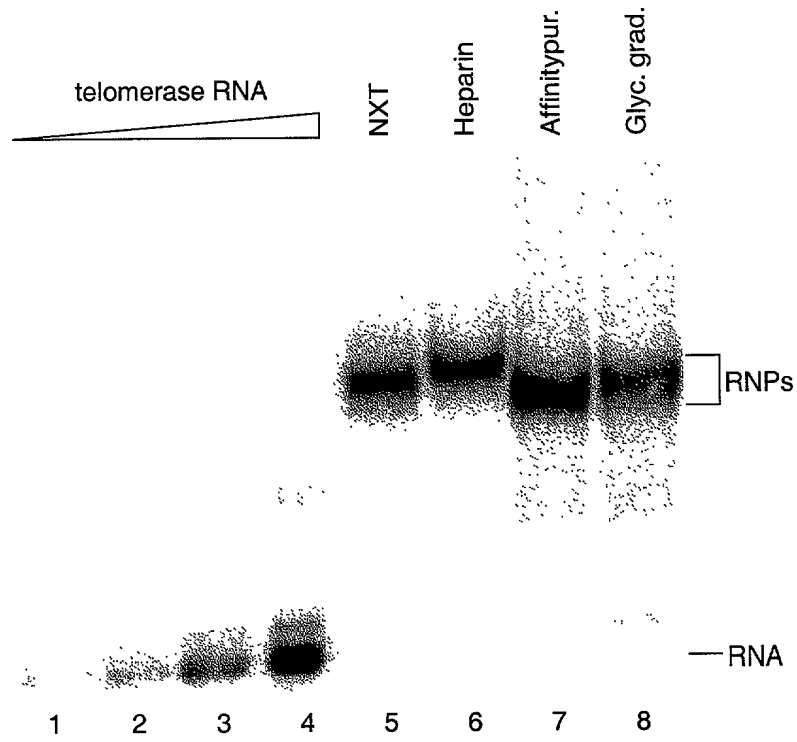


FIG. 2

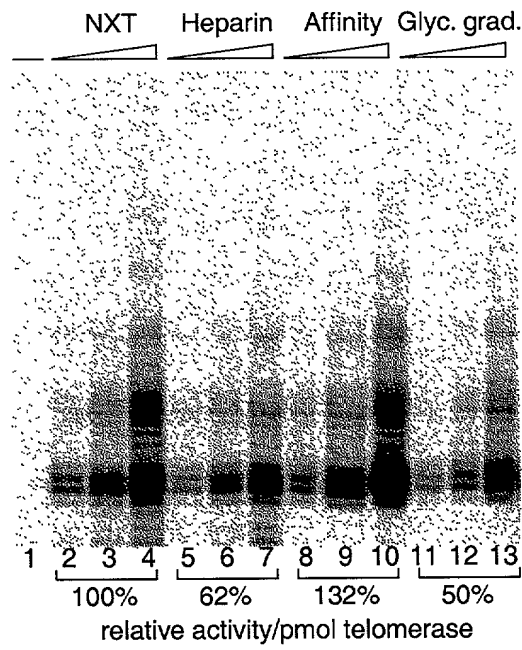


FIG. 3

+

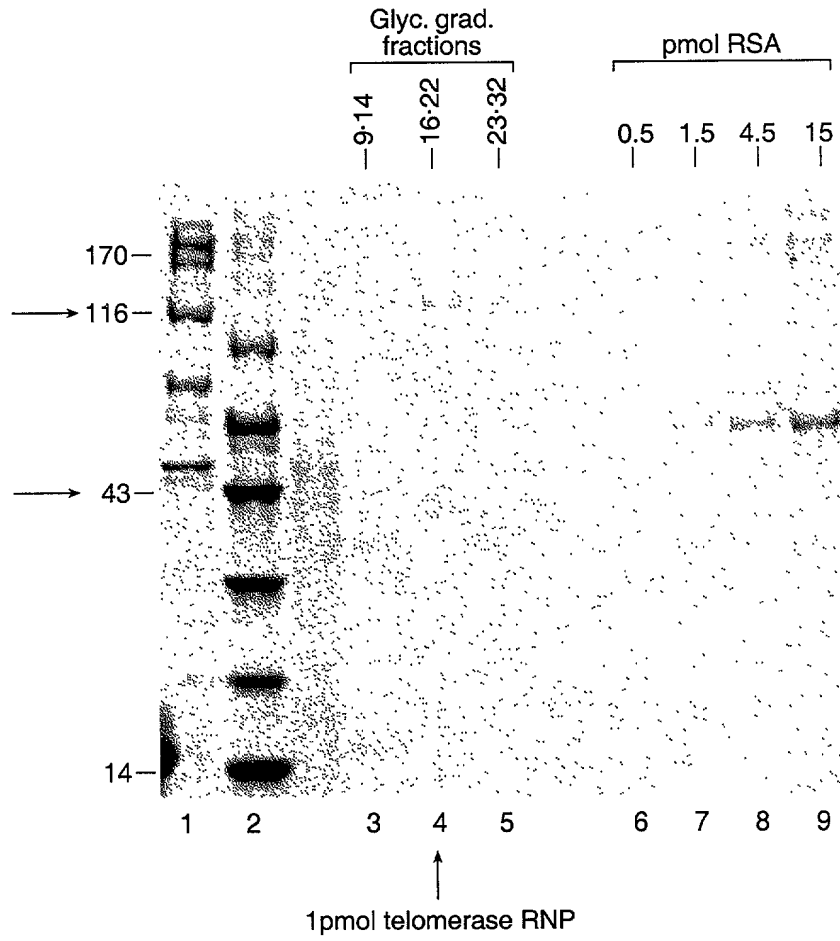


FIG. 4

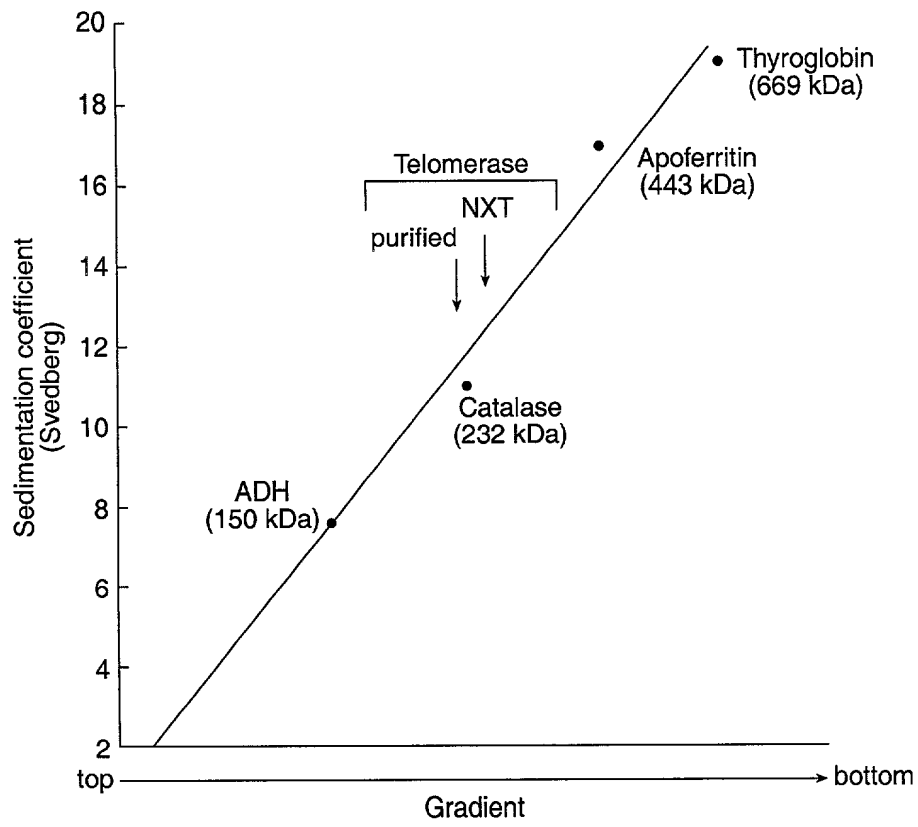


FIG. 5

Telomerase:

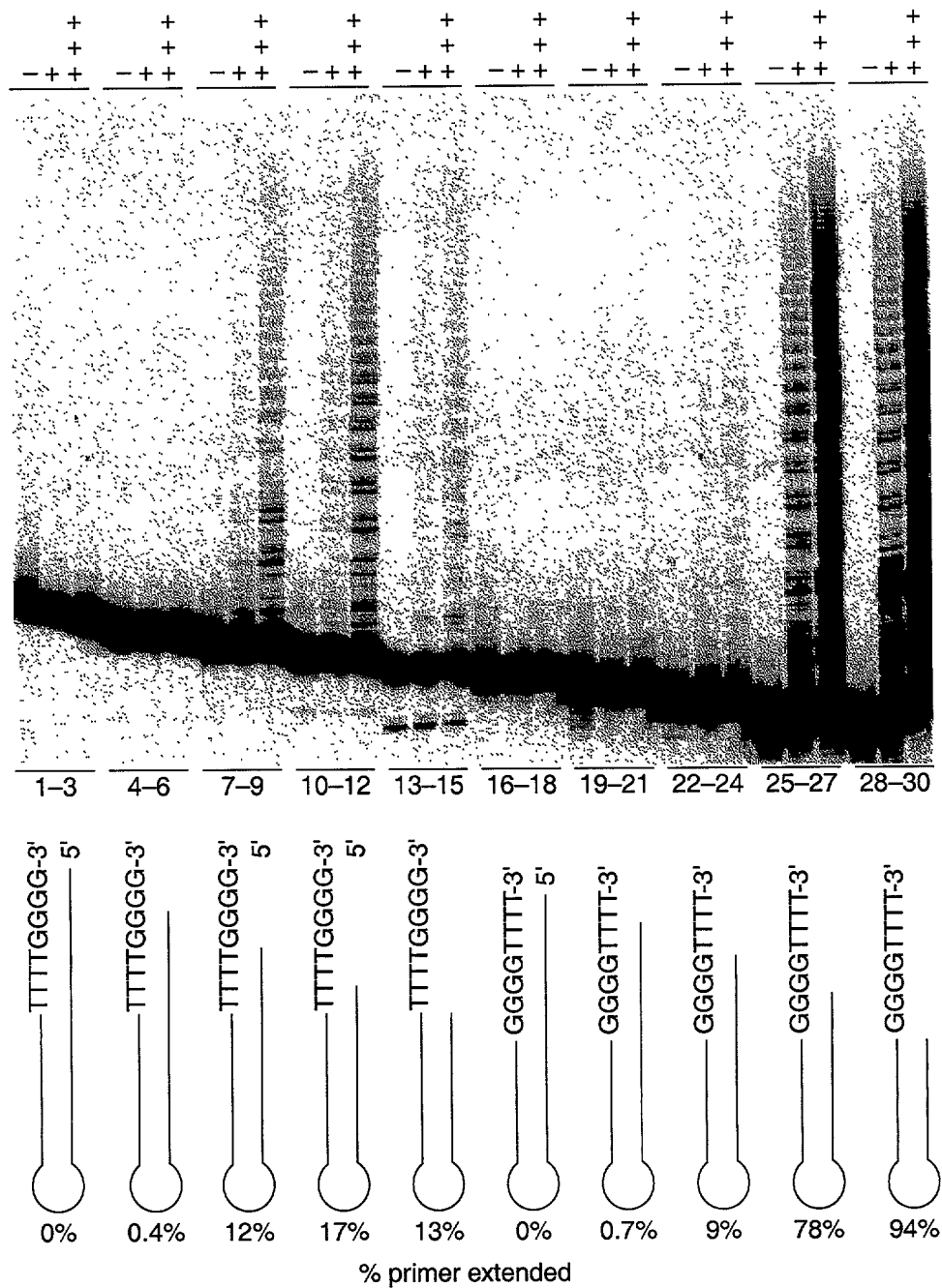


FIG. 6

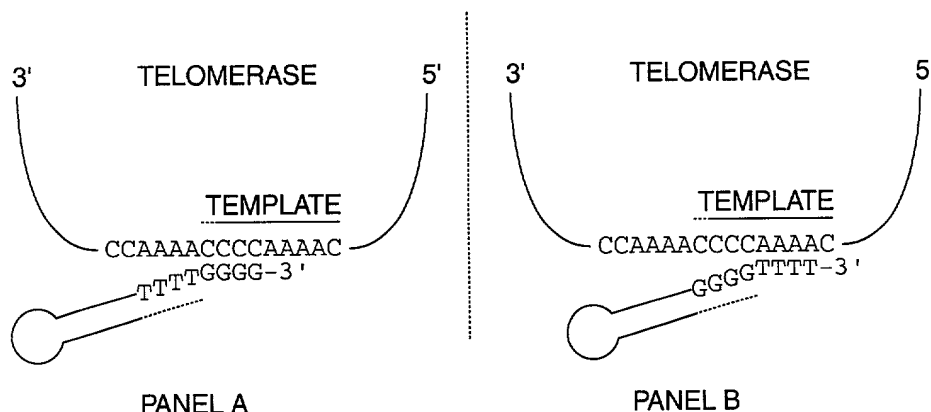


FIG. 7

1	CCCCAAAACC	CCAAAACCCC	AAAACCCCTA	TAAAAAAGA	AAAAATTGAG
51	GTAGTTTAGA	AATAAAATAT	TATTCCTGCA	CAATGGAGA	TGGATATTGA
101	TTTGGATGAT	ATAGAAAATT	TACTTCCTAA	TACATTC AAC	AAGTATAGCA
151	GCTCTGTAG	TGACAAGAAA	GGATGCAAAA	CATTGAAATC	TGGCTCGAAA
201	TCGCCCTCAT	TGACTATTCC	AAAGTTGCAA	AAACAATTAG	AGTTCTACTT
251	CTCGGATGCA	AATCTTTATA	ACGATTCTTT	CTTGAGAAAA	TTAGTTTTAA
301	AAAGCGGAGA	GCAAAGAGTA	GAAATTGAAA	CATTACTAAT	GTTTAAATAA
351	AATCAGGTAA	TGAGGATTAT	TCTATTTTTT	AGATCACTTC	TTAAGGAGCA
401	TTATGGAGAA	AATTACTTAA	TACTAAAAGG	TAAACAGTTT	GGATTATTTT
451	CCTAGCCAAC	AATGATGAGT	ATATTTAAAT	CATATGAGAA	TGAGTCAAAG
501	GATCTCGATA	CATCAGACTT	ACCAAAGACA	AACTCGCTAT	AAAACGCAAG
551	AAAAAGTTTG	ATAATCGAAC	AGCAGAAGAA	CTTATTGCAT	TTACTATTTCG
601	TATGGGTTTT	ATTACAATTG	TTTTAGGTAT	CGACGGTGAA	CTCCCGAGTC
651	TTGAGACAAT	TGAAAAAGCT	GTTTACAAC	GAAGGAATCG	CAGTTCTGAA
701	AGTTCTGATG	TGTATGCCAT	TATTTTGTGA	ATTAATCTCA	AATATCTTAT
751	CTCAATTTAA	TGGATAGCTA	TAGAAACAAA	CCAAATAAAC	CATGCAAGTT
801	TAATGGAATA	TACGTTAAAT	CCTTTGGGAC	AAATGCACAC	TGAATTTATA
851	TTGGATTCTT	AAAGCATAGA	TACACAGAAT	GCTTTAGAGA	CTGATTTAGC
901	TTACAACAGA	TTACCTGTTT	TGATTACTCT	TGCTCATCTC	TTATATCTTT
951	AAAAGAAGCA	GGCGAAATGA	AAAGAAGACT	AAAGAAAGAG	ATTTCAAAAT
1001	TTGTTGATTC	TTCTGTAACC	GGAATTAACA	ACAAGAATAT	TAGCAACGAA
1051	AAAGAAGAAG	AGCTATCACA	ATCCTGATTC	TTAAAGATTT	CAAAAATTCC
1101	AGGTAAGAGA	GATACATTCA	TTAAAATTCA	TATATTATAG	TTTTTCATTT
1151	CACAGCTGTT	ATTTTCTTTT	ATCTTAACAA	TATTTTGTGA	TTAGCTGGAA
1201	GTAAAAAGTA	TCAAATAAGA	GAAGCGCTAG	ACTGAGGTAA	CTTAGCTTAT
1251	TCACATTCAT	AGATCGACCT	TCATATATCC	AATACGATGA	TAAGGAAACA
1301	GCAGTCATCC	GTTTTAAAAA	TAGTGCTATG	AGGACTAAAT	TTTTAGAGTC
1351	AAGAAATGGA	GCCGAAATCT	TAATCAAAAA	GAATTGCGTC	GATATTGCAA
1401	AAGAATCGAA	CTCTAAATCT	TTCGTTAATA	AGTATTACCA	ATCTTGATTG
1451	ATTGAAGAGA	TTGACGAGGC	AACTGCACAG	AAGATCATTA	AAGAAATAAA
1501	GTAACTTTTA	TTAATTAGAG	AATAAACTAA	ATTACTAATA	TAGAGATCAG
1551	CGATCTTCAA	TTGACGAAAT	AAAAGCTGAA	CTAAAGTTAG	ACAATAAAAA
1601	ATACAAACCT	TGGTCAAAAT	ATTGAGGAAG	GAAAAGAAGA	CCAGTTAGCA
1651	AAAGAAAAAA	TAAGGCAATA	AATAAAATGA	GTACAGAAGT	GAAGAAATAA
1701	AAGATTTTAT	TTTTTCAATA	ATTTATTGAA	AAGAGGGGTT	TTGGGGTTTT
1751	GGGGTTTTTG	GG			

FIG. 11

+

7/59

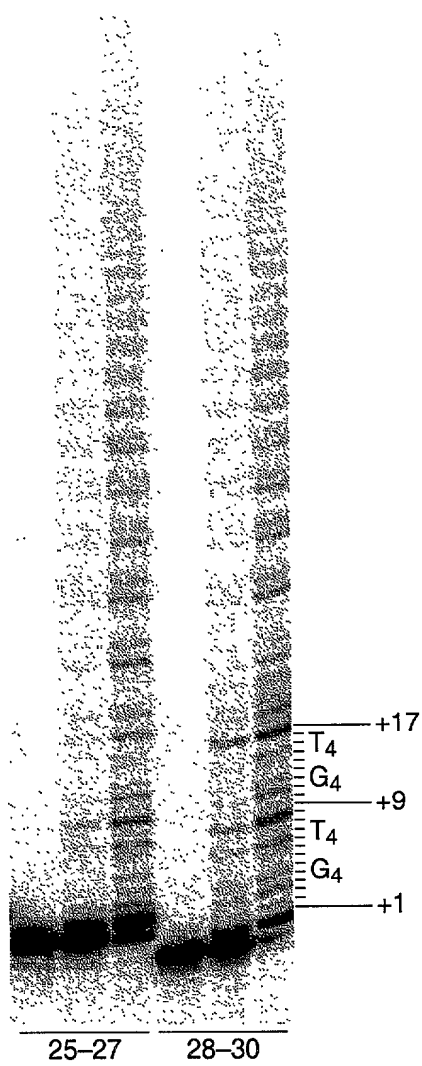


FIG. 8

+

1 AAAACCCCAA AACCCCAAAA CCCCTTTT TAG AGCCCTGCAG TTGGAAATAT  
 51 AACCTCAGTA TTAATAAGCT CAGATTTTAA ATATTAATTA CAAAACCTAA  
 101 ATGGAGGTTG ATGTTGATAA TCAAGCTGAT AATCATGGCA TTCACTCAGC  
 151 TCTTAAGACT TGTGAAGAAA TTAAAGAAGC TAAAACGTTG TACTCTTGGA  
 201 TCCAGAAAGT TATTAGATGA AGAAATCAAT CTCAAAGTCA TTATAAAGAT  
 251 TTAGAAGATA TTAAAATATT TGCGCAGACA AATATTGTTG CTACTCCACG  
 301 AGACTATAAT GAAGAAGATT TTAAAGTTAT TGCAAGAAAA GAAGTATTTT  
 351 CAACTGGACT AATGATCGAA CTTATTGACA AATGCCTTAGT TGAACCTCTT  
 401 TCATCAAGCG ATGTTTCAGA TAGACAAAAA CTTCAATGAT TTGGATTTC A  
 451 ACTTAAGGGA AATCAATTAG CAAAGACCCA TTTATTAACA GCTCTTTTCAA  
 501 CTCAAAGCA GTATTTCTTT CAAGACGAAT GGAACCAAGT TAGAGCAATG  
 551 ATTGGAAATG AGCTCTTCCG ACATCTCTAC ACTAAATATT TAATATTCCA  
 601 GCGCAACTTCT GAAGGAATC TTGTTCAATT TTGCGGGAAT AACGTTTTTTG  
 651 ATCATTTGAA AGTCAACGAT AAGTTTGACA AAAAGCAAAA AGGTGGAGCA  
 701 GCAGACATGA ATGAACCTCG ATGTTGATCA ACCTGCAAAAT ACAATGTCAA  
 751 GAATGAGAAA GATCACTTTC TCAACAACAT CAACGTGCCG AATTGGAATA  
 801 ATATGAAATC AAGAACCAGA ATATTTTATT GCACTCATTT TAATAGAAAT  
 851 AACCAATTCT TCAAAAAGCA TGAGTTTGTG AGTAACAAAA ACAATATTTT  
 901 AGCGATGGAC AGAGCTCAGA CGATATTAC GAATATATTC AGATTTAATA  
 951 GAATTAGAAA GAAGCTAAAA GATAAGGTTA TCGAAAAAAT TGCCTACATG  
 1001 CTTGAGAAAG TCAAAGATTT TAACTTCAAC TACTATTTAA CAAAATCTTG  
 1051 TCCTCTTCCA GAAAATTGGC GGAACGGAA ACAAAAATC GAAAACCTGA  
 1101 TAAATAAAAC TAGAGAAGAA AAGTCGAAGT ACTATGAAGA GCTGTTTAGC  
 1151 TACACAAC TGATAAATG CGTCACACAA TTTATTAATG AATTTTCTA  
 1201 CAATATACTC CCCAAAGACT TTTTGACTGG AAGAAACCGT AAGAATTTTC  
 1251 AAAAGAAAAGT TAAGAAATAT GTGGAACATA ACAAGCATGA ACTCATTCAC  
 1301 AAAAATTTAT TGCTTGAGAA GATCAATACA AGAGAAATAT CATGGATGCA  
 1351 GGTTGAGACC TCTGCAAAGC ATTTTATTA TTTTGATCAC GAAAACATCT  
 1401 ACGTCTTATG GAAATTGCTC CGATGGATAT TCGAGGATCT CGTCGTCTCG  
 1451 CTGATTAGAT GATTTTCTTA TGTCACCGAG CAACAGAAAA GTTACTCCAA  
 1501 AACCTATTAC TACAGAAAGA ATATTGGA CGTCATTATG AAAATGTCAA  
 1551 TCGCAGACTT AAAGAAGGAA ACGCTTGCTG AGGTCCAAGA AAAAGAGGTT  
 1601 GAAGAATGGA AAAAGTCGCT TGGATTGCA CCTGGAAC TCAGACTAAT  
 1651 ACCGAAGAAA ACTACTTTC GTCCAAATAT GACTTTC AAT AAGAAGATTG  
 1701 TAAATTGAGA CCGGAAGACT ACAAAATTA CTACAAATAC GAAGTTATTG  
 1751 AACTCTCACT TAATGCTTAA GACATTGAAG AATAGAATGT TTAAAGATCC  
 1801 TTTTGGATTG GCTGTTTTTA ACTATGATGA TGTAATGAAA AAGTATGAGG  
 1851 AGTTTGTTTG CAAATGGAAG CAAGTTGGAC AACC AAAACT CTCTTTGCA  
 1901 ACTATGGATA TCGAAAAGTG ATATGATAGT GTAAACAGAG AAAA ACTATC  
 1951 AACATTCCCTA AAAACTACTA AATTACTTTC TTCAGATTTC TGGATTATGA  
 2001 CTGCACAAAT TCTAAAGAGA AAGAATAACA TAGTTATCGA TTCGAAAAAC  
 2051 TTTAGAAAGA AAGAAATGAA AGATTATTTT AGACAGAAAT TCCAGAAGAT  
 2101 TGCATTGAA GGAGGACAAT ATCCAACCTT ATTCAGTGT CTGAAAATG  
 2151 AACAAAATGA CTTAAATGCA AAGAAAACAT TAATTGTTGA AGCAAAGCAA  
 2201 AGAAATTATT TTAAGAAAGA TAACTTACTT CAACCAGTCA TTAATATTTG  
 2251 CCAATATAAT TACATTAAC TTAATGGGAA GTTTTATAAA CAAACAAAAG  
 2301 GAATTCCTCA AGGTCTTTGA GTTTCATCAA TTTTGTCATC ATTTTATTAT  
 2351 GCAACATTAG AGGAAAGCTC CTTAGGATTC CTTAGAGATG AATCAATGAA

FIG. 9



2401 CCCTGAAAAT CCAAATGTTA ATCTTCTAAT GAGACTTACA GATGACTATC  
 2451 TTTTGATTAC AACTCAAGAG AATAATGCAG TATTGTTTAT TGAGAAACTT  
 2501 ATAAACGTAA GTCGTGAAAA TGGATTTAAA TTCAATATGA AGAAACTACA  
 2551 GACTAGTTTT CCATTAAGTC CAAGCAAATT TGCAAAATAC GGAATGGATA  
 2601 GTGTTGAGGA GCAAAATATT GTTCAAGATT ACTGCGATTG GATTGGCATC  
 2651 TCAATTGATA TGAAAACCTCT TGCTTTAATG CCAAATATTA ACTTGAGAAT  
 2701 AGAAGGAATT CTGTGTACAC TCAATCTAAA CATGCAAACA AAGAAAGCAT  
 2751 CAATGTGGCT CAAGAAGAAA CTAAAGTCGT TTTTAATGAA TAACATTACC  
 2801 CATTATTTTA GAAAGACGAT TACAACCGAA GACTTTGCGA ATAAACTCT  
 2851 CAACAAGTTA TTTATATCAG GCGGTTACAA ATACATGCAA TGAGCCAAAG  
 2901 AATACAAGGA CCACTTTAAG AAGAACTTAG CTATGAGCAG TATGATCGAC  
 2951 TTAGAGGTAT CTAAAATTAT ATACTCTGTA ACCAGAGCAT TCTTTAAATA  
 3001 CCTTGTTGTC AATATTAAGG ATACAATTTT TGGAGAGGAG CATTATCCAG  
 3051 ACTTTTTCTT TAGCACACTG AAGCACTTTA TTGAAATATT CAGCACAAAA  
 3101 AAGTACATTT TCAACAGAGT TTGCATGATC CTCAAGGCAA AAGAAGCAAA  
 3151 GCTAAAAAGT GACCAATGTC AATCTCTAAT TCAATATGAT GCATAGTCGA  
 3201 CTATTCTAAC TTATTTTGA AAGTTAATT TCAATTTTTG TCTTATATAC  
 3251 TGGGGTTTTG GGGTTTTGG GTTTTGGGG

FIG. 9  
(CONTINUED)

1 MEVDVDNQAD NHGIHSALKT CEEIKEAKTL YSWIQKVIRC RNQSQSHYKD  
 51 LEDIKIFAQT NIVATPRDYN EEDFKVIARK EVFSTGLMIE LIDKCLVELL  
 101 SSSDVSDRQK LQCFGFQLKG NQLAKTHLLT ALSTQKQYFF QDEWNQVRAM  
 151 IGNELEFRHLY TKYLIFQRTS EGTLVQFCGN NVFDHLKVND KFDKKQKGGG  
 201 ADMNEPRCCS TCKYNVKNK DHFLNNINVP NWNMKSRTTR IFYCTHFNRN  
 251 NQFFKKHEFV SNKNNISAMD RAQTIFTNIF RFNRIRKKLK DKVIEKIAM  
 301 LEKVKDFNFN YYLTKSCPLP ENWRERKQKI ENLINKTREE KSKYYEELFS  
 351 YTTDNKCVTQ FINEFFYNIL PKDFLTGRNR KNFQKKVKKY VELNKHელი  
 401 KNLLLEKINT REISWMQVET SAKHFYYFDH ENIYVLWKL RWIFEDLVVS  
 451 LIRCFYVTE QQKSYSKTY YRKNWVIM KMSIADLKKE TLAEVQEKEV  
 501 EEWKSLGFA PGKLRLIPK TTFRPIMTFN KKIVNSDRKT TKLTNTKLL  
 551 NSHMLKTLK NRMFKDPFGF AVFNYDDVMK KYEEFVCKWK QVGQPKLFFA  
 601 TMDIEKCYDS VNREKLSTFL KTTKLLSSDF WIMTAQILKR KNNIVIDSKN  
 651 FRKKEMKDYF RQKFQKIALE GGQYPTLFSV LENEQNDLNA KKTLLIVEAKQ  
 701 RNYFKKDNLL QPVINICQYN YINFNGKFYK QTKGIPQGLC VSSILSSFY  
 751 ATLESSLGF LRDESMNPEN PNVNLLMRLT DDYLLITTQE NNAVLFIKEL  
 801 INVSRENGFK FNMKKLQTSF PLSPSKFAKY GMDSVEEQNI VQDYCDWIGI  
 851 SIDMKTLALM PNINLRIEGL LCTLNLMQT KKASMWLKKK LKSFLMNNIT  
 901 HYFRKTITTE DFANKTLNKL FISGGYKYM CAKEYKDHFK KNLAMSSMID  
 951 LEVSKIIYSV TRAFFKYLVC NIKDTIFGEE HYPDFFLSTL KHFIelfSTK  
 1001 KYIFNRVCM LKAKEAKLKS DQCQSLIQYD A

FIG. 10

CCCCCAAACCCCAAAACCCCAAAACCCCTATAAAAAAGAAAAATTGAGGTAGTTTAGA  
 1 -----+-----+-----+-----+-----+ 60  
 GGGGTTTTGGGGTTTTGGGGTATTTTTTCTTTTAACTCCATCAAATCT  
  
 a P Q N P K T P K P L \* K K K K L R \* F R -  
 b P K T P K P Q N P Y K K R K N C G S L E -  
 c P K P Q N P K T P I K K E K I E V V \* K -  
  
 AATAAAATATTATCCCGCACAAATGGAGATGGATATTGATTGGATGATATAGAAAATT  
 61 -----+-----+-----+-----+-----+ 120  
 TTATTTTATAATAAGGGCGTGTTCCTCTACCTATAACTAAACCTACTATATCTTTTAA  
  
 a N K I L F P H K W R W I L I W M I \* K I -  
 b I K Y Y S R T N G D G Y C F G C Y R K F -  
 c \* N I I P A Q M E M D I D L D D I E N L -  
  
 TACTTCTAATACATTCAACAAGTATAGCAGCTCTTGTAGTGACAAGAAAGGATGCAAAA  
 121 -----+-----+-----+-----+-----+ 180  
 ATGAAGGATTATGTAAGTTGTTTCATATCGTCGAGAACATCACTGTTCTTTCTACGTTTT  
  
 a Y F L I H S T S I A A L V V T R K D A K -  
 b T S \* Y I Q Q V \* Q L L \* \* Q E R M Q N -  
 c L P N T F N K Y S S S C S D K K G C K T -  
  
 CATTGAAATCTGGCTCGAAATCGCCTTCATTGACTATTCCAAAGTTGCAAAAACAATTAG  
 181 -----+-----+-----+-----+-----+ 240  
 GTAACTTTAGACCGAGCTTTAGCGGAAGTAAGTATAAGGTTTCAACGTTTTTGTAAATC  
  
 a H C N L A R N R L H C L F Q S C K N N \* -  
 b I E I W L E I A F I D Y S K V A K T I R -  
 c L K S G S K S P S L T I P K L Q K Q L E -  
  
 AGTTCTACTTCTCGGATGCAAAATCTTTATAACGATTCTTTCTTGAGAAAATTAGTTTTAA  
 241 -----+-----+-----+-----+-----+ 300  
 TCAAGATGAAGAGCCTACGTTTAGAAATATTGCTAAGAAAGAACTCTTTTAATCAAAATT  
  
 a S S T S R M Q I F I T I L S C E N \* F \* -  
 b V L L L G C K S L \* R F F L E K I S F K -  
 c F Y F S D A N L Y N D S F L R K L V L K -  
  
 AAAGCGGAGAGCAAAGAGTAGAAATTGAAACATTACTAATGTTTAAATAAAATCAGGTAA  
 301 -----+-----+-----+-----+-----+ 360  
 TTTCGCCTCTCGTTTCTCATCTTTAACTTTGTAATGATTACAAATTTATTTTAGTCCATT  
  
 a K A E S K E \* K L K H Y \* C L N K I R \* -  
 b K R R A K S R N C N I T N V \* I K S G N -  
 c S G E Q R V E I E T L L M F K \* N Q V M -  
  
 TGAGGATTATTCTATTTTTTAGATCACTTCTTAAGGAGCATTATGGAGAAAATTACTTAA  
 361 -----+-----+-----+-----+-----+ 420  
 ACTCCTAATAAGATAAAAAATCTAGTGAAGAATTCCTCGTAATACCTCTTTTAATGAATT  
  
 a C G L F Y F L D H F L R S I M E K I T \* -  
 b E D Y S I F \* I T S \* G A L W R K L L N -  
 c R I I L F F R S L L K E H Y G E N Y L I -

FIG. 12

```

TACTAAAAGGTAAACAGTTTGGATTATTTCCCTAGCCAACAATGATGAGTATATTTAAATT
421 -----+-----+-----+-----+-----+-----+-----+ 480
ATGATTTTCCATTTGTCAAACCTAATAAAGGGATCGGTTGTTACTACTCATATAATTTAA

a Y * K V N S L D Y F P S Q Q C C V Y * I -
b T K R * T V W I I S L A N N D E Y I K F -
c L K G K Q F G L F P * P T M M S I L N S -

CATATGAGAATGAGTCAAAGGATCTCGATACATCAGACTTACCAAAGACAAACTCGCTAT
481 -----+-----+-----+-----+-----+-----+ 540
GTATACTCTTACTCAGTTTCTAGAGCTATGTAGTCTGAATGGTTTCTGTTTGAGCGATA

a H M R M S Q R I S I H Q T Y Q R Q T R Y -
b I C E C V K G S R Y I R L T K D K L A I -
c Y E N E S K D L D T S D L P K T N S L * -

AAAAACGCAAGAAAAAGTTTGATAATCGAACAGCAGAAGAACTTATTGCATTTACTATTCG
541 -----+-----+-----+-----+-----+-----+ 600
TTTTGCGTTCTTTTCAAACCTATTAGCTTGTCTTCTGAATAACGTAAATGATAAGC

a K T Q E K V C * S N S R R T Y C I Y Y S -
b K R K K K F D N R T A E E L I A F T I R -
c N A R K S L I I E Q Q K N L L H L L F V -

TATGGGTTTTATTACAATTGTTTGTAGGTATCGACGGTGAACCTCCCGAGTCTTGAGACAAT
601 -----+-----+-----+-----+-----+-----+ 660
ATACCCAAAATAATGTAAACAAATCCATAGCTGCCACTTGAGGGCTCAGAACTCTGTTA

a Y G F Y Y N C F R Y R R C T P E S C D N -
b M G F I T I V L G I D G E L P S L E T I -
c W V L L Q L F * V S T V N S R V L R Q L -

TGAAAAAGCTGTTTACAACCTGAAGGAATCGCAGTTCTGAAAGTTCTGATGTGTATGCCAT
661 -----+-----+-----+-----+-----+-----+ 720
ACTTTTTTCGACAAATGTTGACTTCCTTAGCGTCAAGACTTTCAAGACTACACATACGGTA

a C K S C L Q L K E S Q F C K F * C V C H -
b E K A V Y N C R N R S S E S S D V Y A I -
c K K L F T T E G I A V L K V L M C M P L -

TATTTTGTGAATTAATCTCAAATATCTTATCTCAATTTAATGGATAGCTATAGAAACAA
721 -----+-----+-----+-----+-----+-----+ 780
ATAAAACACTTAATTAGAGTTTATAGAATAGAGTTAAATTACCTATCGATATCTTTGTTT

a Y F V N * S Q I S Y L N L M D S Y R N K -
b I L C I N L K Y L I S I * W I A I E T N -
c F C E L I S N I L S Q F N G * L * K Q T -

CCAAATAAACCATGCAAGTTTAAATGGAATATACGTTAAATCCTTTGGGACAAATGCACAC
781 -----+-----+-----+-----+-----+-----+ 840
GGTTTATTTGGTACGTTCAAATTACCTTATATGCAATTTAGGAAACCCTGTTTACGTGTG

a P N K P C K F N G I Y V K S F G T N A H -
b Q I N H A S L M E Y T L N P L G Q M H T -
c K * T M Q V * W N I R * I L W D K C T L -

TGAATTTATATTGGATTCTTAAAGCATAGATACACAGAATGCTTTAGAGACTGATTTAGC
841 -----+-----+-----+-----+-----+-----+ 900
ACTTAAATATAACCTAAGAATTTTCGTATCTATGTGTCTTACGAAATCTCTGACTAAATCG

a C I Y I G F L K H R Y T E C F R D C F S -
b E F I L D S * S I D T Q N A L E T D L A -
c N L Y W I L K A * I H R M L * R L I * L -

```

FIG. 12  
(CONTINUED)

TTACAACAGATTACCTGTTTTGATTACTCTTGCTCATCTCTTATATCTTTAAAAGAAGCA  
 901 -----+-----+-----+-----+-----+ 960  
 AATGTTGTCTAATGGACAAAACCTAATGAGAACGAGTAGAGAATATAGAAATTTCTTCGT  
  
 a L Q Q I T C F D Y S C S S L I S L K E A -  
 b Y N R L P V L I T L A H L L Y L \* K K Q -  
 c T T D Y L F C L L L L I S Y I F K R S R -  
  
 GGCGAAATGAAAAGAAGACTAAAGAAAGAGATTTCAAAATTTGTTGATTCTTCTGTAACC  
 961 -----+-----+-----+-----+-----+ 1020  
 CCGCTTTACTTTTCTTCTGATTTCTTTCTCTAAAGTTTAAACAACCTAAGAAGACATTGG  
  
 a G E M K R R L K K E I S K F V D S S V T -  
 b A K C K E D \* R K R F Q N L L I L L \* P -  
 c R N E K K T K E R D F K I C C F F C N R -  
  
 GGAATTAACAACAAGAATATTAGCAACGAAAAAGAAGAGCTATCACAATCCTGATTC  
 1021 -----+-----+-----+-----+-----+ 1080  
 CCTTAATTGTTGTTCTTATAATCGTTGCTTTTTCTTCTCTCGATAGTGTAGGACTAAG  
  
 a G I N N K N I S N E K E E E L S Q S C F -  
 b E L T T R I L A T K K K K S Y H N P D S -  
 c N \* Q Q E Y \* Q R K R R R A I T I L I L -  
  
 TTAAAGATTTCAAAAATTCAGGTAAGAGAGATACATTCAATAAAATTCATATATTATAG  
 1081 -----+-----+-----+-----+-----+ 1140  
 AATTTCTAAAGTTTTTAAGGTCCATTCTCTCTATGTAAGTAATTTTAAGTATATAATATC  
  
 a L K I S K I P G K R D T F I K I H I L \* -  
 b \* R F Q K F Q V R E I H S L K F I Y Y S -  
 c K D F K N S R \* E R Y I H \* N S Y I I V -  
  
 TTTTTCATTTACAGCTGTTATTTTCTTTTATCTTAACAATATTTTTTGATTAGCTGGAA  
 1141 -----+-----+-----+-----+-----+ 1200  
 AAAAAGTAAAGTTCGACAATAAAAGAAAATAGAATTGTTATAAAAACTAATCGACCTT  
  
 a F F I S Q L L F S F I L T I F F D \* L E -  
 b F S F H S C Y F L L S \* Q Y F L I S W K -  
 c F H F T A V I F F Y L N N I F C L A G S -  
  
 GTAAAAAGTATCAAATAAGAGAAGCGCTAGACTGAGGTAACCTAGCTTATTCACATTCAT  
 1201 -----+-----+-----+-----+-----+ 1260  
 CATTTTTTCATAGTTTATTCTCTTCGCGATCTGACTCCATTGAATCGAATAAGTGAAGTA  
  
 a V K S I K \* E K R \* T E V T \* L I H I H -  
 b \* K V S N K R S A R L R \* L S L F T F I -  
 c K K Y Q I R E A L D C G N L A Y S H S \* -  
  
 AGATCGACCTTCATATATCCAATACGATGATAAGGAAACAGCAGTCATCCGTTTTAAAAA  
 1261 -----+-----+-----+-----+-----+ 1320  
 TCTAGCTGGAAGTATATAGGTTATGCTACTATTCCTTTGTCGTCAGTAGGCAAAATTTTT  
  
 a R S T F I Y P I R C \* G N S S H P F \* K -  
 b D R P S Y I Q Y D D K E T A V I R F K N -  
 c I D L H I S N T M I R K Q Q S S V L K I -  
  
 TAGTGCTATGAGGACTAAATTTTTAGAGTCAAGAAATGGAGCCGAAATCTTAATCAAAAA  
 1321 -----+-----+-----+-----+-----+ 1380  
 ATCAGGATACTCCTGATTTAAAAATCTCAGTTCTTTACCTCGGCTTTAGAATTAGTTTTT  
  
 a \* C Y E D \* I F R V K K W S R N L N Q K -  
 b S A M R T K F L E S R N G A E I L I K K -  
 c V L C G L N F \* S Q E M E P K S \* S K R -

FIG. 12  
(CONTINUED)

GAATTGCGTCGATATTGCAAAAGAATCGAACTCTAAATCTTTCGTTAATAAGTATTACCA  
 1381 -----+-----+-----+-----+-----+-----+ 1440  
 CTTAACGCAGCTATAACGTTTCTTAGCTTGAGATTTAGAAAGCAATTATTCATAATGGT

a E L R R Y C K R I E L \* I F R \* \* V L P -  
 b N C V D I A K E S N S K S F V N K Y Y Q -  
 c I A S I L Q K N R T L N L S L I S I T N -

ATCTTGATTGATTGAAGAGATTGACGAGGCAACTGCACAGAAGATCATTAAAGAAATAAA  
 1441 -----+-----+-----+-----+-----+ 1500  
 TAGAACTAACTAACTTCTCTAACTGCTCCGTTGACGTGTCTTCTAGTAATTTCTTTATTT

a I L I D C R D \* R G N C T E D H \* R N K -  
 b S C L I E E I D E A T A Q K I I K E I K -  
 c L D C L K R L T R Q L H R R S L K K \* S -

GTAACCTTTTATTAATTAGAGAATAAACTAAATTACTAATATAGAGATCAGCGATCTTCAA  
 1501 -----+-----+-----+-----+-----+ 1560  
 CATTGAAAATAATTAATCTCTTATTTGATTTAATGATTATATCTCTAGTCGCTAGAAGTT

a V T F I N \* R I N \* I T N I E I S D L Q -  
 b \* L L L I R E \* T K L L I \* R S A I F N -  
 c N F Y \* L E N K L N Y \* Y R D Q R S S I -

TTGACGAAATAAAAGCTGAACTAAAGTTAGACAATAAAAAATACAAACCTTGGTCAAAAT  
 1561 -----+-----+-----+-----+-----+ 1620  
 AACTGCTTTATTTTCGACTTGATTTCATCTGTATTTTTTATGTTTGAACCAAGTTT

a L T K \* K L N \* S \* T I K N T N L G Q N -  
 b C R N K S C T K V R Q \* K I Q T L V K I -  
 c D E I K A E L K L D N K K Y K P W S K Y -

ATTGAGGAAGGAAAAGAAGACCAGTTAGCAAAAGAAAAATAAGGCAATAAATAAAATGA  
 1621 -----+-----+-----+-----+-----+ 1680  
 TAACTCCTTCCTTTTCTTCTGGTCAATCGTTTCTTTTTTATCCGTTATTTATTTTACT

a I E E G K E D Q L A K E K I R Q \* I K C -  
 b L R K E K K T S \* Q K K K \* G N K \* N E -  
 c C G R K R R P V S K R K N K A I N K M S -

GTACAGAAAGTGAAGAAATAAAAGATTTATTTTTTTCAATAATTTATTGAAAAGAGGGGTT  
 1681 -----+-----+-----+-----+-----+ 1740  
 CATGCTTCACTTCTTTATTTTCTAAATAAAAAAGTTATTAAATAACTTTTCTCCCCAA

a V Q K C R N K R F I F F N N L L K R G V -  
 b Y R S E E I K D L F F S I I Y C K E G F -  
 c T E V K K \* K I Y F F Q \* F I E K R G F -

TTGGGGTTTTGGGGTTTGGGG  
 1741 -----+-----+-----+-----+ 1762  
 AACCCCAAACCCCAAACCC

a L G F W G F G -  
 b W G F G V L G -  
 c G V L G F W -

FIG. 12  
 (CONTINUED)

2 EVDVQNQADNHGIHSALKTCEEIKEAKTLYSWIQKVIRCRNQSQSHYKDL 51  
 |:::| ::::|: :| :::| ::: . | |..|..|  
 19 ELELEMQENQNDIQVRVK...IDDPKQY..LVNVTAACLLQEGSYQDK 62  
 52 EDIKIFAQTNIVATPRDYNEEDFKVIARKEVF.STGLMIELIDKCLVELL 100  
 :: :::: :.|. |..|. |: :::| | ::::|: :  
 63 DERRYIITKALL...EVAESDPEFICQLAVYIRNELYIRTTTNYIVAF. 107  
 101 SSSDVSDRQKLQCFGQLKGNQLAKTHLLTALSTQKQYFFQDEWNQVRAM 150  
 ::: :::: :. |: :::: |::| :  
 108 .....CVVHKNTQPFIEKYFNKAVLLPNDLLEVCEFAQVLYI 144  
 151 IGNELFRHLYTKYLIFQRTSEGLVQFCGNNVFDHLKVNDKFDKKQKGA 200  
 ::: | :|| | : :::: ..|... ::: :  
 145 FDATEFNKLY.....LDRILSQDIRKELTFRKCLQRCVRSKF 181  
 201 ADMNE...PRCCSTCKYNVKNKEDHFLNNINVPNWNMKSRTTRIFYCTHF 247  
 :::|| :::|: :. |... :|.. |..|: |:  
 182 SEFNEYQLGKYCTES..QRKKTMFYRLSVTNKQKWDQTKK..... 220  
 248 NRNNQFFKKHEFVSNKNNISAMDRAQTIFTNIFRFRNRIRKCLKDKVIEKI 297  
 |.. :..| : :::: | : : :|: ::: | |..|: ||  
 221 .RKENLLTKLQAIKESDKSKRETG.....DIMNVEDAIKALKPAVMKKI 264  
 298 AYMLEKVKDFNFNYLTKSCPLPENWRERKQKIENLINKTREEKSKYYEE 347  
 |. . :| : : : : :| : : :|..| : : :  
 265 AKRQNAMK.....KHMKAPKIPNSTLESKYLTFKD 294  
 348 LFSYTTDNKCVTQFINEFFYNILPKDFTGRNRKNFQKKVKYVELNKHE 397  
 |:: :.. :. |..|..|: : : : : : : :|..|  
 295 LIKFCHISEP.....KERVYKILGKKYPKTEEEYKAAFGDSASAPFN.PE 338  
 398 LIHKNLLEKINTREISWMQVETSAKHFFYYFDHENIYVLWKLRLWIFEDL 447  
 | : : :| : : : : : : : : : :| : : :  
 339 LAGRMKIEISKWENELSAKGNTAEVWDNLISSNQLPYMAMLRNLSN.. 386  
 448 VVSLIRCFYVTEQQKSYSKTYYYRKNIWVIMKMSIADLKKETLAEVQE 497  
 : : :| : : :|  
 387 .....ILKAGVSD..... 394  
 498 KEVEEWKSLGFAFGKRLRLIPKKTTFRPIMTFNKKIVNSDRKTTKLTTNT 547  
 : : :| : : :|  
 395 .....TTHS 398  
 548 KLLNSHMLMLKTLKNRMFKDPFGFAVFNYDDVMKKYEEFVCKWKVQVGPKL 597  
 | : : :| : : :| : : :| : : :|  
 399 IVINK.....ICEPKAVENSKM 415  
 598 FFATMDIEKCYDSVNREKLSTFLKTTKLSSDFWIMTAQILKRKNNIVID 647  
 | : : : :| : : : :| : : : :| : : : :| : : :  
 416 F..PLQFFSAIEAVN.EAVTKGFKA...RENMNKLGQIEAVKE..VVE 457  
 648 SKNFRKKEMKDYFRQKFQKIALEGGQYPTLFSVLENEQNLDLNAKKTIVE 697  
 : : :| : : :| : : :| : : :| : : :| : : :| : : :  
 458 KTDEEKKDM.....ELEQTEEGEFVKVNEGIGKQYINSIELAIK 496  
 698 AKQRNYFKDNLQPVINICQYNYINFNGKFYKQTKGIPQGLCVSSILSS 747  
 :  
 497 IAVNKNLDEIKGHTAIFSDVSGSMSTSMSSGAKKYSVRTCLECALVLGL 546  
 748 FYYATLEESSLGLRDESMNPENPNVNLMLRLTDDYLLITTQENNAVLFI 797  
 : : :| : : :| : : :| : : :| : : :| : : :| : : :  
 547 MVKQRCEKSSFYIFSSPSSQCNCYLEVDL..... 576

FIG. 13

798 EKLINVSRENGFKFNMKK.LQTSFPLSPSKFAKYGMDSVEEQNIVQDYCD 846  
 577 .....PGDELRPSMQKLLQEKGLGGG..TDFPYECIDWTKNKTTHVD 617  
 847 WIGISIDMKTLALMPNINLRIEGILCTNLNMQTKKASMWLKKKLKSFLM 896  
 618 NIVILSDMMIAEGYSINVRGSSIVNSI.....KKYKDEVN 653  
 897 NNITHYFRKTITTEDFANKTLNKLFISSGGYKYMQCAKEYKD.HFKKNLAM 945  
 654 PNIKIF...AVDLEGYK.....KCLNLGDEFNENNYIKIFGM 687  
 946 SSMIDLEVSKIIYSVTRAFFKYLVCNIKDTIFGEEHYPDFFLSTLKHFI 995  
 688 SDSI.....LKFISAKQGA.....NMVE 706  
 996 IFSTKKYIFNRVC 1008  
 707 VI..KNFALQKIG 717

FIG. 13  
(CONTINUED)

132 LSTQKQYFFQDEWNQVRAMIGNEL.FRHLYTKYLIFQRTSE..GTLVQFC 178  
 1 MSRRNQ.....KKPQAPIGNETNLDVFLQNLVYKYSQIEHYKTQQQQI 43  
 179 GNNVFDHLKVNDKFDKKQKGAADMNEPRCCSTCKYNVKNKDHFLNNIN 228  
 44 KEEDLKLKFKNQDQDGNNGNDDDEE.....NNSNKQQLLRRVN 84  
 229 VPNNWNMKSRTIRFYCTHFNRRNQFFKKHEFVSNNKNISAMDRAQTIFTN 278  
 85 .....QIKQVQLIKK...VGSKEVDLNLNEDENKKN 114  
 279 IFRFNIRKKLKDVKIEKIAYMLEKVKDFNFNYLTKSCPLPENWRERKQ 328  
 115 GLSEQQVKEEQRLRTITEEQVKYQNLVFNMDYQLDLNESGGHRRHRRETDY 164  
 329 KIENLINKTREESKYEEELFSYTTDNKCVTQFINE.FFYNILPKDFLTG 377  
 165 DTEKWFEISHDQK.....NYVSIYANQKTSYCWWLKDYFNK 200  
 378 RNRKNFQKKVKYVELNKHLEIHNLLLEKINTREISWMQVETSAKHFY 427  
 201 NNYDHLNVSINRLE..TEAEFYAFDDFSQTIKLTNNSYQTVNID..... 242  
 428 FDHENIYVLWKLLRWI..FEDLVVSLIRCFYVTEQQKSYSKYTYRKN 475  
 243 VNFDDNLCILALLRFLSLERFNILNIRSSY..TRNQYNFEKIGELLETI 290  
 476 WDVIMKMSIADLKKETLAEVQKEVEEWKSLGFAPGKLRLIPKKTTFRP 525  
 291 FAVVFSHR.....HLQGIHLQVPCEAFQYLVNSSSQISVKDSQLQ 330  
 526 IMTFNKKIVNSDRKTTKLTNTNKLNSHMLKTLKNRMFKDPFGFAVFNY 575  
 331 VYSFSTDCLKLVD..TNKVQDYFKFLQEFPRLTHVSQQAIPVSATNAVENL 378

FIG. 14

FIG. 15

+



```

1 MEMDIDLDDIENL....LPNTFNKYSSSCSDKKGCKTLKSGSKSPS... 42
  |:|. . .|| ..|.. |:| |...: . ||| ..:
491 IELAIIKIAVNKNLDEIKGHTAIFSDVSGSMSTMSGGAKKYGSVRTCLEC 540
43 .LTIPKLQKQ.....LEFYFSDANLYNDSFLRKLVLKSGEQRVEIETLL 85
  |: . : || : :: |... :|::| . : |::| |...: ||
541 ALVLGLMVKORCEKSSFYIFSSPSSQCNCYL.EVDLPGLDELRPMSQKLL 589

```

FIG. 16

```
telomerase p43  LQKQLEFYFSDANLYNDSFLRKVLVLSGGEQRVEIETLLM
human La        ICHQUEYYFYGDFNLPRDKFLKEQI.KLDEGWVPLEIMIK
Xenopus LaA     ICEQIEYYFGDHNLPDKFLKQOI.LLDDGWVPLETMIK
Drosophila La   ILRQVEYYFGDANLNRDKFLREQIGKNEDGWVPLSVLVT
S. c. Lhp1p     CLKQVEFYFSEFNFPYDRFLRTAEK.NDGWVPISTIAT
```

FIG. 18

1	aactcatttta	attactaatt	taatcaacaa	gattgataaa	aagcagtaaa	taaaacccaa
61	tagattttaat	ttagaaaagta	tcaattgaaa	aatggaaatt	gaaaacaact	aagcacaaata
121	gccaaaagcc	gaaaaattgt	ggtgggaact	tgaattagag	atgcaagaaa	accaaaatga
181	tatataagtt	agggttaaga	ttgacgatcc	taagcaatag	ctcgtgaacg	tcactgcagc
241	atgtttgttg	taggaaggta	gttactacta	agataaagat	gaaagaagat	atatcatcac
301	taaaagcactt	cttgagggtg	ctgagttctga	tcctgagttc	atctgctagt	tggcagttcta
361	catccgtaat	gaactttaca	tcagaactac	cactaactac	attgtagcat	tttgtgttgt
421	ccacaagaat	actcaaccat	tcatcgaaaa	gtactttcaac	aaagcagttac	tttgtcctaa
481	tgacttactg	gaagttctgt	aaattgcata	ggtttctctat	atttttgatg	cacttgaatt
541	caaaaatttg	tatcttgata	ggataactttc	ataagatatt	cgtaagggaac	tcactttccg
601	taagtgttta	caaagatgcg	tcagaagcaa	gttttctgaa	ttcaacgaat	actaacttgg
661	taagtattgc	actgaatcct	aacgtaagaa	aacaatgttc	cgttacacct	cagttaccaa
721	caagtaaaaag	tgggattaaa	ctaagaagaa	gagaaaaagag	aatctcttaa	ccaaacttta
781	ggcaataaag	gaatctgaag	ataagtccaa	gagagaaact	ggagacataa	tgaacgttga
841	agtgcaatc	aaggctttaa	aaccagcagt	tatgaagaaa	atagccaaga	gatagaatgc
901	catgaagaaa	cacatgaagg	cacttaaaat	tcctaactct	accttggaat	caaagttact
961	gaccttcaag	gatctcatta	agttctgcca	tattttctgag	cctaaagaaa	gagttctata
1021	gatccttggg	aaaaataacc	ctaagccga	agagaaataac	aaagcagcct	ttgggtgattc
1081	tgcattctgca	cccttcaatc	ctgaattggc	tggaaaagcgt	atgaagattg	aaatctctaa
1141	aacatgggaa	aatgaactca	gtgcaaaagg	caacactgct	gagggtttggg	ataattttaat
1201	ttcaagcaat	taactcccat	atatggccat	gttacgtaac	ttgtctaaaca	tcttaaaagg
1261	cgggtgtttca	gatactacac	actctattgt	gatcaacaag	atttgtgagc	ccaaggccgt
1321	tgagaactcc	aagatgtttc	ctcttcaatt	ctttagtgcc	attgaagctg	ttaatgaagc
1381	agttactaag	ggattcaagg	ccaagaagag	agaaaatatg	aatcttaaag	gtcaaatcga
1441	agcagtaaag	gaagttgttg	aaaaaaccga	tgaagagaag	aaagatatgg	agttggagta
1501	aaccgaagaa	ggagaatttg	ttaaagtcaa	cgaaggaatt	ggcaagcaat	acattaaact
1561	cattgaactt	gcaatcaaga	tagcagttaa	caagaatttta	gatgaaatca	aaggacacac
1621	tgcaatcttc	tctgatgttt	ctggtttctat	gagtaacctca	atgtcagggtg	gagccaagaa
1681	gtatggttcc	gttcgtactt	gtctcgagtg	tgcattagtc	cttggtttga	tggtaaaata
1741	acgttgttgaa	aagtctctat	tctacatctt	cagttcacct	agttctcaat	gcaataagtg
1801	ttacttagaa	gttgatctcc	ctggagacga	actccgtcct	tctatgtaaa	caattttgca
1861	agagaaaagga	aaacttgytg	tgtggtactga	tttccctatg	gagtgcatgt	atgaatggac
1921	aaagaataaaa	actcacgtag	acaatatcgt	tattttgtct	gatattgatga	ttgcagaagg
1981	atattcagat	atcaatgtta	gaggcagttc	cattgttaac	agcatcaaaa	agtacaagga
2041	tgaagtaaat	cctaacatta	aaactctttc	agttgactta	gaaggttacg	gaaagtcgct
2101	taatctaggt	gatgagttca	atgaaaacaa	ctacatcaag	atattcggta	tgaagcattc
2161	aatcttaaag	ttcatttcag	ccaagcaagg	aggagcaaat	atggtcgaag	ttatcaaaaa
2221	ctttgccctt	caaaaaatag	gacaaaagtg	agtttcttga	gattcttcta	taacaaaaat
2281	ctcaccacac	ttttttgttt	tattgcatag	ccattatgaa	atttaaatta	ttatctattt
2341	atttaagtta	cttcatatgt	ttatgtatcg	cagttctatta	gcctattcaa	atgattctgc
2401	aaagaacaaa	aaagattaaa	a			

FIG. 19

+

	Motif A	Motif B
Consensus telomerase p123 Dong (LINE) a1 S.c. (groupII) HIV-RT L8543.12	h--h <b>dh</b> ---h--h GQPKLFFAT <b>MD</b> IEKCYDSV <b>NRE</b> KLSTFLK <b>TT</b> KLL-100-KFYKQ <b>TK</b> GIP <b>QG</b> LCV <b>SS</b> ILSSFYATLEESSLGFL KRN <b>N</b> LHCT <b>VD</b> Y <b>DD</b> YK <b>KAF</b> DSIPHSWLIQVLEI <b>Y</b> KIN-28-RQIAIK <b>KGI</b> Y <b>QG</b> DSL <b>PL</b> WFC <b>LAL</b> N <b>PL</b> SHQ <b>LH</b> ND <b>R</b> FGGS <b>NW</b> FRE <b>VD</b> L <b>KK</b> CFDTIS <b>HD</b> LI <b>IK</b> ELK <b>R</b> YISD-26-HVPV <b>GR</b> VCV <b>QG</b> APT <b>SP</b> ALCN <b>AV</b> LL <b>R</b> LD <b>RL</b> AG <b>LA</b> L <b>KK</b> K <b>S</b> VT <b>VL</b> DVG <b>DA</b> YF <b>SV</b> PL <b>DE</b> DFRKY <b>TAF</b> TP-7-GIRYQ <b>YN</b> VL <b>PQ</b> GW <b>KG</b> SP <b>AI</b> FQSS <b>MT</b> K <b>IL</b> EP <b>FR</b> KQ <b>N</b> VL <b>PE</b> LY <b>FM</b> K <b>FD</b> V <b>K</b> SCYDSIP <b>RNE</b> CM <b>RI</b> LKD <b>AL</b> KN-68-KCY <b>IRE</b> D <b>GL</b> F <b>QG</b> SS <b>LS</b> APIVD <b>LV</b> YDD <b>LL</b> E <b>FY</b> SE <b>FK</b>	h---- <b>QG</b> ---- <b>SP</b>
	Motif C	Motif D
Consensus telomerase p123 Dong (LINE) a1 S.c. (groupII) HIV-RT L8543.12	h--Y <b>h</b> <b>DD</b> h <b>hh</b> -14-L <b>M</b> RL <b>AT</b> <b>DD</b> Y <b>LL</b> IT <b>TO</b> ENN-0-A <b>VL</b> F <b>IE</b> KL <b>IN</b> VS <b>RE</b> NG <b>FK</b> F <b>N</b> M <b>K</b> L <b>Q</b> T-23-QDYCD <b>WI</b> G <b>I</b> S <b>I</b> -16-H <b>L</b> IY <b>MD</b> I <b>K</b> LY <b>AK</b> ND <b>KE</b> -0-M <b>K</b> LID <b>TT</b> TF <b>SND</b> IS <b>MQ</b> F <b>GLD</b> K <b>CK</b> T-25-K <b>CL</b> Y <b>K</b> Y <b>L</b> G <b>F</b> Q <b>Q</b> -55-Y <b>VR</b> Y <b>AD</b> ILIG <b>VL</b> GS <b>KN</b> -2-K <b>IK</b> R <b>DL</b> N <b>N</b> FL <b>NS</b> .L <b>GL</b> T <b>INE</b> E <b>K</b> TL <b>I</b> -4-ET <b>PAR</b> FL <b>E</b> Y <b>NI</b> -4-IYQ <b>Y</b> <b>MD</b> LY <b>VG</b> SH <b>LE</b> IG-1-H <b>RT</b> K <b>IE</b> EL <b>RQ</b> HL <b>LR</b> W <b>GL</b> TT <b>PD</b> K <b>KH</b> Q <b>K</b> -0-EP <b>P</b> FL <b>WM</b> G <b>Y</b> EL -8-IL <b>K</b> L <b>AD</b> <b>DD</b> FL <b>II</b> ST <b>DQ</b> Q.....VINIK <b>K</b> L <b>AM</b> G <b>GF</b> Q <b>KY</b> NA <b>K</b> AN <b>R</b> -41-IR <b>S</b> K <b>S</b> K <b>G</b> I <b>FR</b>	h-h <b>Lgh</b> -h

FIG. 17

+

+

19/59

MEIENNQAQQPKAEKLWWELELEMQENQNDIQVRVKIDDPKQYL  
 VNVTAACLLQEGSYQDKDERRYIITKALLEVAESDPEFICQLA  
 VYIRNELYIRTTTNYIVAFVVKHNTQPFIEKYFNKAVLLPNDL  
 LEVCEFAQVLYIFDATEFKNLYLDRILSQDIRKELTRFKCLQRC  
 VRSKFSEFNEYQLGKYCTESQRKKTMFYRLSVTNKQKWDQTKKK  
 RKENLLTKLQAIKESDKSKRETGDIMNVEDAIKALKPAVMKKI  
 AKRQONAMKKHMKAPKIPNSTLESKYLTFKDLIKFCHISEPKERV  
 YKILGKKYPKTEEEYKAAFSDSASAPFNPPELAGKRMKIEISKW  
 ENELSAKGNTAEVWDNLISSNQLPYMAMLRNLSNLIKAGVSDTT  
 HSIVINKICEPKAVENSKMFPLOFFSAIEAVNEAVTKGFKAKKR  
 ENMNLKGQIEAVKEVVEKTDEEKKDMELEQTEEGEFVKVNEGIG  
 KQYINSIELAIKIAVNKNLDEIKGHTAIFSDVSGSMSTSMSSGA  
 KKYGSVRTCLECALVGLMVKQRCESKFYIFSSPSSQCNCYL  
 EVDLPGDELPSMQKLLQEKGLGGGTDFPYECIDEWTKNKTHV  
 DNIVILSDMMIAEGYSDINVRGSSIVNSIKKYKDEVNPNIKIFA  
 VDLEGYGKCLNLGDEFNENNYIKIFGMSDSILKFISAKQGGANM  
 VEVIKNFALQKIGQK

FIG. 20

MSRRNQKKPQAPIGNETNLDFVLQNLLEVYKSQIEHYKTQQQQIK  
 EEDLKLKFKNQDQDGNNGNDDDEENNSNKQOELLRRVNQIKQ  
 QVQLIKKVGSKVEKDLNLNEDENKKNGLSEQQVKEEQLRITTEE  
 QVKYQNLVFNMDYQLDLNESGGHRRHRETQDYDTEKWFESHQ  
 KNYVSIYANQKTSYCWWLKDYFNKNNDHLNVSINRLETEAEFY  
 AFDDFSQTIKLTNNSYQTVNIDVNFNNLCILALLRFLLSLERF  
 NILNIRSSYTRNQYNFEKIGELLETFIFAVVFSHRHLQGIHLQVP  
 CBAFYQLVNSSSQISVKDSLQVYSFSTDLKLVDTNKVQDYFKF  
 LQEFPRLTHVSQQAIPVSATNAVENLNVLLKKVKHANLNLVSI  
 TQFNDFYFVNLQHLKLEFLEPNILTKQKLENLLSISQSKNL  
 KFLRLNFYTYVAQETSRKQILQATTIKNLKNNKNEETPETKD  
 ETPSESTSGMKFFDHLSELTELEDFSVNLQATQEIYDSLHLLI  
 RSTNLKKFKLSYKEMEKSMDTFIDLKNIYETLNNLKRCVNI  
 SNPHGNIISYELTNKDSTFYKFKLTNLQELQHAKYTFKQNEFQFN  
 NVKSAKIESSLESLEDIDSLCKSIASCKNLQNVNIIASLLYPN  
 NIQKNPFNKPNLLFFKQFEQLKNLENVINCILDQHILNSISEF  
 LEKNKKIKAFILKRYLLQYLDYTKLFTLQQLPELNQVYINQ  
 QLEELTVSEVHKQVWENHKQKAFYEPLCEFTKESQTLQIDFD  
 QNTVSDDSIKKILESISESKYHHYLRNPSQSSSLIKSENEEIQ  
 ELLKACDEKGVLVKAYYKFPLCLPTGTYDYNSDRW

FIG. 22

MKILFEFIQDKLDIDLQTNSTYKENLKC GFHNGLDEILTTCFAL  
 PNSRKIALPCLPGDLSHKAVIDHCIIYLLTGELYNNVLTFGYKI  
 ARNEDVNNSLFCHSANVNVTLKGAAWKMFHSLVGTYYAFVDLLI  
 NYTVIQFNGQFFTQIVGNRCNEPHLPPKWVQSSSSSATAAQIK  
 QLTEPVTNKQFLHKLNNINSSFFPYSKILPSSSSSIKLTDLREA  
 IFPTNLVKIPQRLKVRINLTQKLLKRHKRLNVVSILNSICPPL  
 EGTVDLDSLHLSRQSPKERVLFIIIVILQKLLPQEMFGSKKNKGK  
 IIKNLNLLLSLPLNGYLPFDSLKKLRLKDFRWLFISDIWFTKH  
 NFENLNQLAICFISWLFRLIPKIIQTFYCTEISSTVTIVYFR  
 HDTWNKLITPFIVEYFKTYLVENNVCRNHSYTLNHNHSMRI  
 IPKKSNNEFRIIAIPCRGADEEFTIYKENHKNAIQPTQKILEY  
 LRNKRPTSFTKIYSPTQIADRIKEFKQRLKKFNNVLPPELYFMK  
 FVDKSCYDSIPRMECMRILKDALKNENGFFVRSQYFFNTNTGVL  
 KLFNVVNASRVKPYELYIDNVRTVHLSNQDVINVVEMEIFKTA  
 LWVEDKCYIREDGLFQSSSLAPIVDLVYDDLLEFYSEFKASPS  
 QDTLILKLADDFLIISTDQQQVINIKKLAMGGFQKYNANARDK  
 ILAVSSQSDDDTVIQFCAMHIFVKELEVWKHSSTMNHFHRSKS  
 SKGIFRSLIALFNTRISYKTIDTNLSTNTVLMQIDHVKNISE  
 CYKSAFKDLSINVTQNMQFHSFLQRIIEMTVSGCPITKCDPLIE  
 YEVRFTILNGFLESLSNNTSKFKDNIILLRKEIQHLQAYIYIYI  
 HIVN

FIG. 23

+

0976653 082004

1 tcaatactat taattaataa ataaaaaaaa gcaaaactaca aagaaaatgt caaggcgtaa  
 61 ctaaaaaaaa ccataggctc ctataggcaa tgaacaaaat cttgattttg tattacaaaa  
 121 tctagaagtt tacaaaaagcc agattgagca ttataagacc tagtagtaat agatcaaaga  
 181 ggaggatctc aagcttttaa agttcaaaaa ttaagattag gatggaaact ctggcaacga  
 241 tgatgatgat gaagaaaaca actcaataaa ataataagaa ttattaagga gagtcaatta  
 301 gattaagtag caagtttaat tgataaaaaa agttggttct aaggtagaga aagatttgaa  
 361 tttgaacgaa gatgaaaaca aaaagaatgg actttctgaa tagcaagtga aagaagagta  
 421 attaagaacg attactgaag aataggttaa gtattaaaaa ttagtattta acatggacta  
 481 ccagtttagat ttaaatgaga gtggtggcca tagaagacac agaagagaaa cagattatga  
 541 tactgaaaaa tggtttgaaa tatctcatga ccaaaaaaat tatgtatcaa ttacgccaa  
 601 ctaaaagaca tcatattgtt ggtggcttaa agattatttt aataaaaaa attatgatca  
 661 tcttaatgta agcattaaca gactagaaac tgaagccgaa ttctatgcct ttgatgattt  
 721 ttcacaaaca atcaaaccta ctaataatct ttactagact gttaacatag acgttaattt  
 781 tgataataat ctctgtatct tctgtatgct tagattttta ttatcactag aaagattcaa  
 841 tattttgaaat ataagatctt cttatacaag aaattaatat aattttgaga aaattgggtga  
 901 gctacttgaa actatcttcg cagttgtctt ttctcatcgc cacttacaag gcattcattt  
 961 acaagttcct tgcgaaagcg tctaataatt agttaactcc tcatacataa ttagcggttaa  
 1021 agatagctaa ttatagggtat actctttctc tacagactta aaattagttg acactaacia  
 1081 agtccaagat tatttttaagt tcttataaga attccctcgt ttgactcatg taagctagta  
 1141 ggctatccca gttagtgtca ctaacgctgt agagaacctc aatgttttac ttaaaaagggt  
 1201 caagcatgct aatcttaatt tagtttctat ccctacctaa ttcaattttg atttctactt  
 1261 tgtaatttta taacatttga aattagagtt tggattagaa ccaaatattt tgacaaaaca  
 1321 aaagcttgaa aatcactttt tgagtataaa ataatacaaa aatcttaaat ttttaagatt  
 1381 aaactttttac acctacgttg ctttaagaaac ctccagaaaa cagatattaa aacaagctac  
 1441 aacaatcaaa aatctcaaaa acaataaaaa tcaagaagaa actcctgaaa ctaaaagatga  
 1501 aactccaagc gaaagcacaa gtggtatgaa attttttgat catctttctg aattaacgga  
 1561 gcttcaagat ttcagcgtta acttgtaagc taccacaagaa atttatgata gcttgaccaa  
 1621 acttttgatt agatcaacaa atttaaagaa gttcaaatta agttacaaat atgaaatgga  
 1681 aaagagtaaa atggatacat tcatagatct taagaatatt tatgaaacct taaacaatct  
 1741 taaaagatgc tctgttaata tatcaaatcc tcatggaac atttcttatg aactgacaaa  
 1801 taaagattct actttttata aatttaagct gaccttaaac taagaattat aacacgctaa  
 1861 gtatactttt aagtagaacg aatttttaatt taataacggt aaaagtgcaa aaattgaatc  
 1921 ttctcatta gaaagcttag aagatattga tagtctttgc aaatctattg cttcttgtaa  
 1981 aaattttaca aatgtttaata ttatcgccag tttgctctat cccaacaata tttagaaaaa  
 2041 tccttttcaat aagcccaatc ttctattttt caagcaattt gaataattga aaaatttgga  
 2101 aaatgtatct atcaactgta ttcttgatca gcatatactt aattctattt cagaattctt  
 2161 agaaaagaat aaaaaaataa aagcattcat tttgaaaaga tattatttat tacaatatta  
 2221 tcttgattat actaaattat ttaaaacact tcaatagtta cctgaattaa attaaagttta  
 2281 cattaattag caattagaag aattgactgt gagtgaagta cataagtaag tatgggaaaa  
 2341 ccacaagcaa aaagctttct atgaaccatt atgtgagttt atcaaagaat catcctaaac  
 2401 ccttttagcta atagattttg accaaaacac tgtaagtgat gactctatta aaaagatttt  
 2461 agaattctata tctgagtcta agtatcatca ttatttgaga ttgaacccta gttaatctag  
 2521 cagtttaatt aaatctgaaa acgaagaaat ttaagaactt ctcaaagctt gcgacgaaaa  
 2581 aggtgtttta gtaaaagcat actataaatt ccctctatgt ttaccaactg gtacttatta  
 2641 cgattacaat tcagatagat ggtgattaat taaatattag tttaaataaa tattaatat  
 2701 tgaatatttc tttgcttatt atttgaataa tacatacaat agtcattttt agtgttttga  
 2761 atatatttta gttatttaatt tcattatttt aagtaataaa ttatttttca atcatttttt  
 2821 aaaaaatcg

FIG. 21

+

21/59

Oxytricha  
Euplotes

LCVSYILSSFYYANLEENALQFLRKESMDPEKPETNLLMRLT  
LCVSSILSSFYYATLESSLGFLRDESMNPENPNVNLMLRLT

FIG. 24

ATTTATACTCATGAAAATCTTATTCGAGTTCATTCAAGACAAGCTTGACATTGATCTACA  
GACCAACAGTACTTACAAAGAAAATTTAAAATGTGGTCACTTCAATGGCCTCGATGAAAT  
TCTAACTACGTGTTTCGCACTACCAAATTCAGAAAAATAGCATTACCATGCCTTCCTGG  
TGACTTAAGCCACAAAGCAGTCATTGATCACTGCATCATTTACCTGTTGACGGGCGAATT  
ATACAACAACGTACTAACATTTGGCTATAAAATAGCTAGAAATGAAGATGTCAACAATAG  
TCTTTTTTGGCATTCTGCAAATGTTAACGTTACGTTACTGAAAGGCGCTGCTTGGAAAAAT  
GTTCCACAGTTTGGTCGGTACATACGCATTTCGTTGATTTATTGATCAATTATACAGTAAT  
TCAATTTAATGGGCAGTTTTTCACTCAAATCGTGGGTAACAGATGTAACGAACCTCATCT  
GCCGCCCAAATGGGTCCAACGATCATCTCATCATCCGCAACTGCTGCGCAAATCAAACA  
ACTTACAGAACCAGTGACAAATAAACAATTCTTACACAAGCTCAATATAAATTCCTCTTC  
TTTTTTTCCCTTATAGCAAGATCCTTTCCTTCATCATCATCTATCAAAAAGCTAACTGACTT  
GAGAGAAGCTATTTTCCCAAAATTTGGTTAAAATTCCTCAGAGACTAAAGGTACGAAT  
TAAATTTGACGCTGCAAAAGCTATTAAAGAGACATAAGCGTTTGAATTACGTTTCTATTTT  
GAATAGTATTTGCCCCACCATTGGAAGGGACCGTATTGGACTTGTCGCATTTGAGTAGGCA  
ATCACCAAAGGAACGAGTCTTGAAATTTATCATTGTTATTTTACAGAAGTTATTACCCCA  
AGAAATGTTTGGCTCAAAGAAAAATAAAGGAAAAATATCAAGAATCTAAATCTTTTATT  
AAGTTTACCTTAAATGGCTATTTACCATTGTAGTGTGTTGAAAAAGTTAAGATTAAA  
GGATTTTCGGTGGTTGTTTCATTTCTGATATTTGGTTCACCAAGCACAAATTTGAAAACCT  
GAATCAATTTGGCGATTTGTTTCATTTCTGGCTATTTAGACAACATAATCCCAAAATAT  
ACAGACTTTTTTTTACTGCACCGAAATATCTTCTACAGTGACAATTGTTTACTTTAGACA  
TGATACCTGGAATAAACTTATCACCCCTTTTATCGTAGAATATTTTAAGACGTACTTAGT  
CGAAAACAACGTATGTAGAAACCATAATAGTTACACGTTGTCCAATTTCAATCATAGCAA  
AATGAGGATTATACCAAAAAAAGTAATAATGAGTTTCAGGATTATTGCCATCCCATGCAG  
AGGGGCAGACGAAGAAGAAATTCACAATTTATAAGGAGAATCACAAAAATGCTATCCAGCC  
CACTCAAAAAATTTTGAATACCTAAGAAACAAAAAGCCGACTAGTTTACTAAAAATATA  
TTCTCCAACGCAAAATAGCTGACCGTATCAAAAGAAATTTAAGCAGAGACTTTTAAAGAAAT  
TAATAATGTCTTACCAGAGCTTTATTTTCATGAAATTTGATGTCAAATCTTGCTATGATTC  
CATACCAAGGATGGAATGTATGAGGATACTCAAGGATGCGCTAAAAAATGAAAATGGGTT  
TTTCGTTAGATCTCAATATTTCTTCAATACCAATACAGGTGTATTGAAGTTATTTAATGT  
TGTTAACGCTAGCAGAGTACCAAAACCTTATGAGCTATACATAGATAATGTGAGGACGGT  
TCATTTATCAAATCAGGATGTTATAAACGTTGTAGAGATGGAAATATTTAAAACAGCTTT  
GTGGGTTGAAGATAAGTGCTACATTAGAGAAGATGGTCTTTTTTCAGGGCTCTAGTTTATC  
TGCTCCGATCGTTGATTTGGTGTATGACGATCTTCTGGAGTTTATAGCGAGTTTAAAGC  
CAGTCCTAGCCAGGACACATTAATTTTAAACTGGCTGACGATTTCTTTATAATATCAAC  
AGACCAACAGCAAGTGATCAATATCAAAAAGCTTGCCATGGGCGGATTTCAAAAAATATAA  
TGCGAAAGCCAATAGAGACAAAATTTTAGCCGTAAGCTCCCAATCAGATGATGATACGGT  
TATTCAATTTTGTGCAATGCACATATTTGTTAAAGAATTGGAAGTTTGGAAACATTCAAG  
CACAATGAATAATTTCCATATCCGTTTCGAAATCTAGTAAAGGGATATTTTCGAAGTTAAT  
AGCGCTGTTTAAACTAGAAATCTCTTATAAAACAATTGACACAAATTTAAATTCACAAA  
CACCGTTCTCATGCAAATTGATCATGTTGTAAAGAACATTTTCGGAATGTTATAAATCTGC  
TTTTAAGGATCTATCAATTAATGTTACGCAAAATATGCAATTCATTCGTTCTTACAACG  
CATCATTGAAATGACAGTCAGCGGTTGTCCAATTACGAAATGTGATCCTTTAATCGAGTA  
TGAGGTACGATTACCATATTGAATGGATTTTGGAAAGCCTATCTTCAAACACATCAAA  
ATTTAAAGATAATATCATTCCTTTTGAGAAAGGAAATTCACACTTGCAAGC

FIG. 26

+

097653 08204  
100280 " 0526450

+

human		
tez1		
EST2		
p123		
	Motif 0	
	AKFLHMLMSVYVVELLRSSFYVVTETTFQKNR	
	ISEIEWLVLGKRSNAKMCCLDSDFEKKRQIFAEFIYWLXNSFIIPILQSFYITESSDLRNR	
	LKDFRWLFISD---IWFTKHNFENLNQLAICFISWLFRLPKIIQTFYCYCTEISSVT-	
	TREISWMQVET-SAKHFYFFDHEN-IYVLWKLRLRWIFEDLVVSLIRCFYFVTEQQKSYSK	
	*** **	
	Motif 1	
human		
tez1		
EST2		
p123		
	LFFYRKSVWSKLSQIGIRQHLKRVQLRDVSEAEVRQHREARPALLTSRLRFPKP--DGL	
	TVYFRKDIWKLRCRPF-I-TSMKWEAFEKINENNVRMDTQK-TTLPPAVIRLLPKK--NTF	
	IVYFRHDTWKNLLITPFIYVYFKTYLVENNVCRNHSYTLS--NFNHSMRRIIPKKSNNEF	
	TYYYRKNINWDVIMKMSI-ADLKKETLAEVQKEVEEWWKS-LGFAPGKRLRLPKK--TTF	
	*** **	
	Motif 2	
human		
tez1		
EST2		
p123		
	RPIVNM DYVVGARTFRREKRAERLT SRVKALF-SVLNYERA	
	RLITN-LRKRFLIKMGSKKMLVSTNQTLRPVASILKHLINEESSGIPFNLEVYMKLLTF	
	RIIAIPCRGADEEFTIYKENHNKNAIQPTQKILEYLRNKRPTSTTKIYSPQTQIADRIKEF	
	RPIMTFNKKIVNSDRKTTTKLTNTTKLLNSHLMKTLKN-RMFKDPFGFAVFNYYDDVMKKY	
	* *	
	Motif 3 (A)	
tez1		
EST2		
p123		
	KKDLLKHRMFGR-KKYFVFRIDIKSCYDRIKQDLMFRIVKK-KLKDPEFVIRKYATIHATS	
	KQRLKKFNVLPELYFMKFDVKSCYDSIPRMCEMRILKD-ALKNENGFFVRSQYFFENTN	
	EEFVCKWKQVGQPKLFFAATMDIEKCYDSVNREKLSTFLKTTKLLSSDFWIMTAQIILKRKN	
	* *	

FIG. 25

+

+

23/59

AKFLHWLMSVYVVELLRSFFYVTETTFQKNRLFFYRKSVWSKLQSIGIRQHLKR  
VQLRDVSEAEVRQHREARPALLTSRLRFIPKPDGLRPVNM DYVVGARTFRREK  
RAERLTSRVKALFSVLNYERA

FIG. 27

GCCAAGTTCCTGCACTGGCTGATGAGTGTGTACGTCGTCGAGCTGCTCAGGTC  
TTTCTTTTATGTCACGGAGACCACGTTTCAAAAGAACAGGCTCTTTTCTACC  
GGAAGAGTGTCTGGAGCAAGTTGCAAAGCATTGAATCAGACAGCACATTGAAG  
AGGGTGCAGCTGCGGGACGTGTCGGAAGCAGAGGTGAGGCAGCATCGGGAAGC  
CAGGCCCGCCCTGCTGACGTCCAGACTCCGCTTCATCCCCAAGCCTGACGGGC  
TGCGGCCGATTGTGAACATGGACTACGTCGTGGGAGCCAGAACGTTCCGCAGA  
GAAAAGAGGGCCGAGCGTCTCACCTCGAGGGTGAAGGCACTGTTACGCGTGT  
CAACTACGAGCGGGCGCG

FIG. 28

MTEHHTPKSRILRFLENQYVYLC TLNDYVQLVLRGSPASSYSNICERLRS DVQTSFSIFLHSTVVG F  
DSKPDEGVQFSSPKCSQSELIANVVQMFDSEFERRNLLMKGFSMNHEDFRAMHVNGVQNDLVSTF  
PNYLISILESKNWQLLLEIIGSDAMHYLLSKGSIFEALPNDNYLQISGIPLFKNNVFEETVSKRRKR  
TIETSITQNK SARKEVSWNSISISRFSIFYRSSYKKFKQDLYFNLHSICDRNTVHMLQWIFPRQFG  
LINAQVQQLHKV IPLVSQSTVVPKRLLKVYPLIEQTAKRLHRISLSKVYNHYCPYIDTHDDEKILS  
YSLKPNQVFAFLRSILVRVFPKLIWGNQRIFEIILKDLETFLKLSRYESFSLHYLMSNIKISEIEWL  
VLGKRSNAKMCLSDFEK RKQIFAEFIYWLYNSFIIPI LQSFFYITESSDLRNRTVYFRKDIWKLLCR  
PFITSMKMEAFEKINENNVRMDTQKTTLP PAVIRLLPKKNTFRLITNL RKRFLIKMGSNKKMLVSTN  
QTLRPVASILKHLINEESSGIPFNLEVYMKLLTFKKDLLKHRMFGRKKYFVRIDIKSCYDRIKQDLM  
FRIVKKKLDPEFVIRKYATIHATSDRATKNFVSEAFSYFDMVPFEKVQLLSMKTSDTLFVDFVDY  
WTKSSSEIFKMLKEHLSGHIVKIGNSQYLQKVGIPQGSILSSFLCHFYMEDLIDEYLSFTKKKGSVL  
LRVDDDFLFI TVNKKDAKKFLNLSLRGF EKHNFTSLEKTVINFENSNGIINNTFFNESKKRMPFFG  
FSVNMRSLDTLLACPKIDEALFNSTSVELTKHMGKSFFYKILRSSLASFAQVFIDITHNSKFNSCCN  
IYRLGYSMCRAQAYLKRMDIFIPQRMFITDLLNVIGRKIWKKLAEILGYTSRRFLSSAEVKWLFC  
LGMRDGLKPSFKYHPCFEQLIYQFQSLTDLIKPLRPVLRQVFLHRRRIAD

FIG. 29

+

097663 0204

ggtaccgatttactttcttcttataagtaattggttctctcgaacgctcctctaaatctctggaataatttttacaaga  
actcaataacaataccaagtcaaattccaatatgaagtggttattagtagtgatgataaattttctattttatcgtgctgta  
ccaagtataaggacaaaagaacaacttctctccctcaagaacttttactttataatttacttttccaatatatttgcg  
ggttcgttactttttaaactcggtactgttttagctgctactcttagccaaacgctgtttctaccccgctcatggatat  
agctcttgagtagtcacagaaatccttacaatcttctgtagagactataattagattcattacagtcggtgcatactc  
ttaacatggagccttacactttagatgagtcacgctgcacatgtagtatattggtatcatccaaagtttgccttgaaaag  
gttgataattatttgcaaaatcatgctccttagtggtgtaacccggaagttttttagatgcttgacacgctctagcatg  
attgagatatccaataattctatccactacaactccttaacgggttttttattttctattttctcatctgtgtt  
ccaaatgatcatcgtattaggtcttttccggttttactcctggaaatcgtaaccttttccactattccccctaataga  
ataactaaattagtttcgcttataaattgattagtagtagaagaattggtgattctactcgtgtaattatttagtttaaa  
gatactttgcaaaacattattagctatcattataaaaaaaatccctataataataataataatttgcggtc  
actatttataaaacgttatgatcagtaggacactttgcatatataatagattgcttaattggttacttgaacttgcatt  
gacCGAACACCATACCCCAAAAGCAGGATTCCTCGCTTTCTAGAGAAATCAATATGTATACCTATGTAACCTTAATGATTT  
ATGTACAACCTTGTTTGTAGAGGTCGCCGGAAGCTCGTATAGCAATATATGCAACGCTTGAGAACGCGATGTACAAACG  
TCCTTTTCTATTTTCTCATTCGACTGTAGTCGGCTTCGACAGTAAGCCAGATGAAGGTGTTCAATTTTCTCTCCAAA  
ATGCTCACAGTCAGAGgtatatatttttggtttggatttttctattctcggtatagctaataatg99cagCTAATAGC  
GAATGTTGTAACACAGATGTCGATGAAAGTTTGAGCGTCGAGGAATCTACTGATGAAAGGTTTTCATGGtaaggt  
attctaaattgtgaaataatttacctgcaattactgtttcaaaagagattgtatttaaccgataaaagAATCATGAGATTTTC  
GAGCCATGCATGTAAACGGAGTACAAAATGATCTCGTTTCTACTTTCCTAAATAGTTTATCTATCTACTTACCTGAGTCAAAA  
CAGTGAATGCCATGCATTAATCTATATCCAAAGGAAGTATTTTTCAGGCTCTTCCAAATGACAATACCTTCAGATTTCTG  
GCATACCACCTTTTAAATAATATGTTTGTAGGAACTGTGAGGAACTGTGTCAAAAGAAAGCAACCATTTGTAAGTCACTTA  
CAAAATAAAAGCGCCCGCAAGAGTTTCTGGAATAGCATTTCAATAGTAGTTTAGCATTTTATATTTTAACTTACCTCTATTT  
TAAGAACTTTAAGCAAGgtaaactaaactgttactgttaccctcataactaatttttagtataattttttaaaggatattaccactaaacgatttt  
GTGATCGGAACACAGTACACATGTGGCTTCAATGGATTTTTCCAAGCAATTTGGACTTATAACGCAATTTCAAGTGAAG  
CAATTGCACAAAGTGATTCCTACTGGTATCACAGAGTACAGTTGTCCTCCAAACGCTCTCTAAAGGTATACCTTTAATTGA  
ACAAACAGCAAGCGACTCCATCGTATTTCTCTATCAAAAGTTTACAAACCATTTATGCCCATATATGACACCCACCGATG  
ATGAAAATTCCTTAGTTATTCCTTAAAGCGAACCGAGGTGTTTGGCTTCTTCGATCCATCTTGTTCGAGTGTTCCT  
AAATTAATCTGGGGTAACCAAGGATATTTGAGATAATATTAAGgtattgtataaaatttattaccactaaacgatttt  
accagACCTCGAAACTTTCTTGAATAATATCGAGATACGAGTCTTTTAGTTTACATTTTACATTTAATGAGTAAATAAGgttaa  
tatgccaaaattttttaccattaaatcaaatcagATTTTCAGAAATGAAATGGCTAGTCTCTTGGAAAAAGGTCAAAATGCG  
AAAAATGTCITTAAGTGATTTTGTAGAAACGCAAGCAAAATATTTGCGAATTCATCTACTGGCTATACAAATCGTTTATAAT  
ACCTATTTTACAACTTTTCTTTTATATCACTGAATCAAGTATTTACGAAATCGAACTGTTTATTTTAGAAAAAGATATTT  
GGAAACTCTTGTGCCGACCCCTTTTATATACATCAATGAAAAATGGAAGGTTTGAAAAAATAAACGAGgtatttttaaagttatt  
ttttgcaaaaagctaaataatttccagAACAAATGTTAGGATGGATCTCAGAAAACTACTTTGCTCCAGCAGTTATTCGT  
TATTACCTAAGAGAAATACCTTTTCGTCATATTACGAATTAAGAAAAAGATTTCTTAATAAGgtatttaaatttttggtcat  
caatgtacttttacttctaactatttagcagATGGGTTCAACAAAAAATGTTAGTCAGTACGAAACCAACTTTACG  
ACCTGTGGCATCGATACGAAACATTTAATCAATGAAGAAAGTAGTGTATTCATTTAACTTGGAGGTTTACATGAAG

FIG. 30



**FIG. 30**  
**(CONTINUED)**

+

EST2 pep	FFYCTEISST	VTIVYFRHDT	WN----	KLIT	P-----	FIVE	YFK-TYLVEN	40
Euplotes pep	FFYVTEQQKS	YSKTYYYRKN	IWDVI-MKMS	IAD----	LKK	ETLA--	EVQE	43
Trans of tetrahymen	-----KHKE	GSQIFYRKP	IWKLVSKLTI	VKVRIQFSEK	NKQMKNNFYQ			44
Consensus	FFY.TE..K.	.S..YYRK.	IW....	KL..	....	F..K	.....V..	50
EST2 pep	NVCRNHNSY-	-----	TLSNFNHSM	RUIPKKSNNE	RUIPKKSNNE			79
Euplotes pep	KEVEEWKKS	-----	GFAPGKG	RUIPKKTT--	FRPTMTFNKK			78
Trans of tetrahymen	KIQLEENLE	KVEEKLIPED	SFQKYPQGKL	RUIPKKGS--	FRPTMTFLRK			92
Consensus	K...E.....	-----	....F..GKL	RUIPKK...--	FRPTMTF.RK			100
EST2 pep	ADDEEFTIYK	ENHKNAIQPT	QKILEYLRNK	RPTSFTKIYS	PTQIADRIKE			129
Euplotes pep	IVNSDRKTTK	LTTNTKLLNS	HMLKTLKN-	-----	RMFK	-DPFGFAVFN		120
Trans of tetrahymen	DKQNIK---	LNLNQILMDS	QLVFRNLKD-	-----	ML-G	-QKIGYSVFD		130
Consensus	.....K..K	LN.N..L..S	QL.L..LKN-	-----	....	---IG..VF.		150
EST2 pep	FKORLLKKFN	NVL-----	PELYFMKFD	VKSCYD				157
Euplotes pep	YD-DVMKKIYE	EFVCKWKQVH	CPKLEFFATMD	IEKCYD				155
Trans of tetrahymen	NK-QISEKFA	QFTEKWNKG	RPCLYYYVTL-	-----				158
Consensus	.K-....KKF.	.F..KWK..G	.P..LYF.T.D	...CYD				186

FIG. 31

+

A-1: AKF LHW LMS VYV VEL LRS FFY VTE TTF Q  
A-2: LFF YRK SVW SKL QSI GIR QHL KRV QLR DVS  
A-3: PAL LTS RLR FIP KPD GLR PIV NMD YVV

[illegible]

$\frac{1017}{4}$   
 $\begin{array}{cccccccc} & & t & & t & & c & \\ & t & a & & a & & g & & c & & c & & t & & c & & g \\ 5' - & cag & acc & aaa & gga & att & cca & taa & gg & -3' \\ & Q & T & K & G & I & P & Q & G \end{array}$

5 (c')

FIG. 34

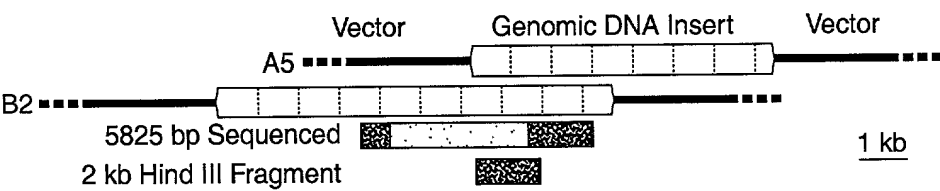


FIG. 33A

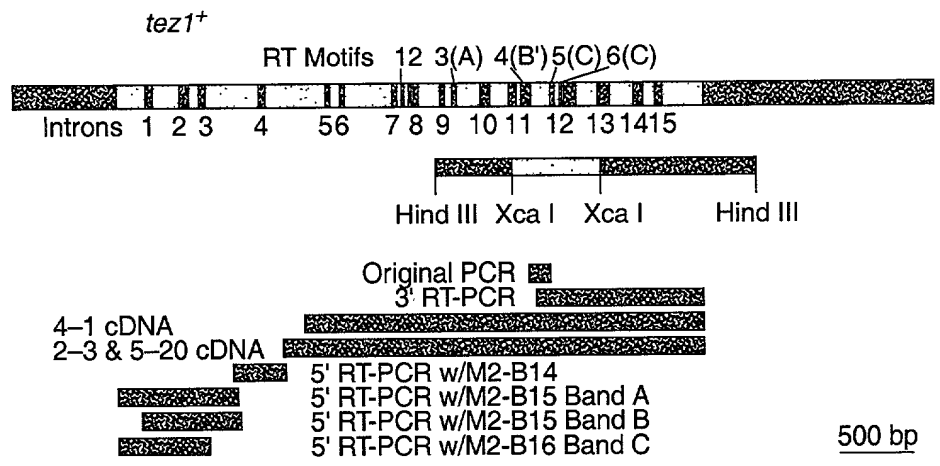


FIG. 33B

+

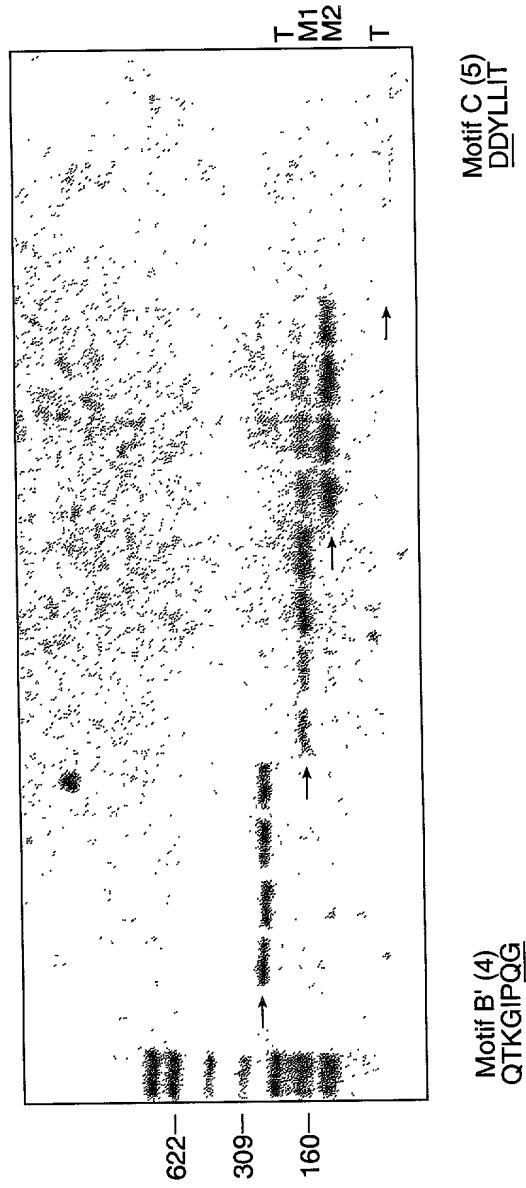


FIG. 35

+

**FIG. 36**

+

GAA GAT TTG ATT GAT GAA TAC CTA TCG TTT ACG AAA AAG AAA GGA TCA GTG TTG TTA CGA  
 CTT CTA AAC TAA CTA CTT ATG GAT AGC AAA TGC TTT TTC CCT AGT CAC AAC AAT GCT  
 E D L I D E Y L S F T K K G S V L L R

GTA GTC gac gac tac ctc ctc atc acc  
 CAT CAG ctg ctg atg gag gag tag tgg

V V D D Y L L I T

<---- ctg ctg atg gag gag tag tgg  
 a a a a a a a  
 t t t t t t t  
 c c c c c c c  
Poly 1

.....gac gat ttc ctc ttt ata aca..... <---Actual Genomic Sequence  
 D D F L F I T

FIG. 36  
 (CONTINUED)

+

+

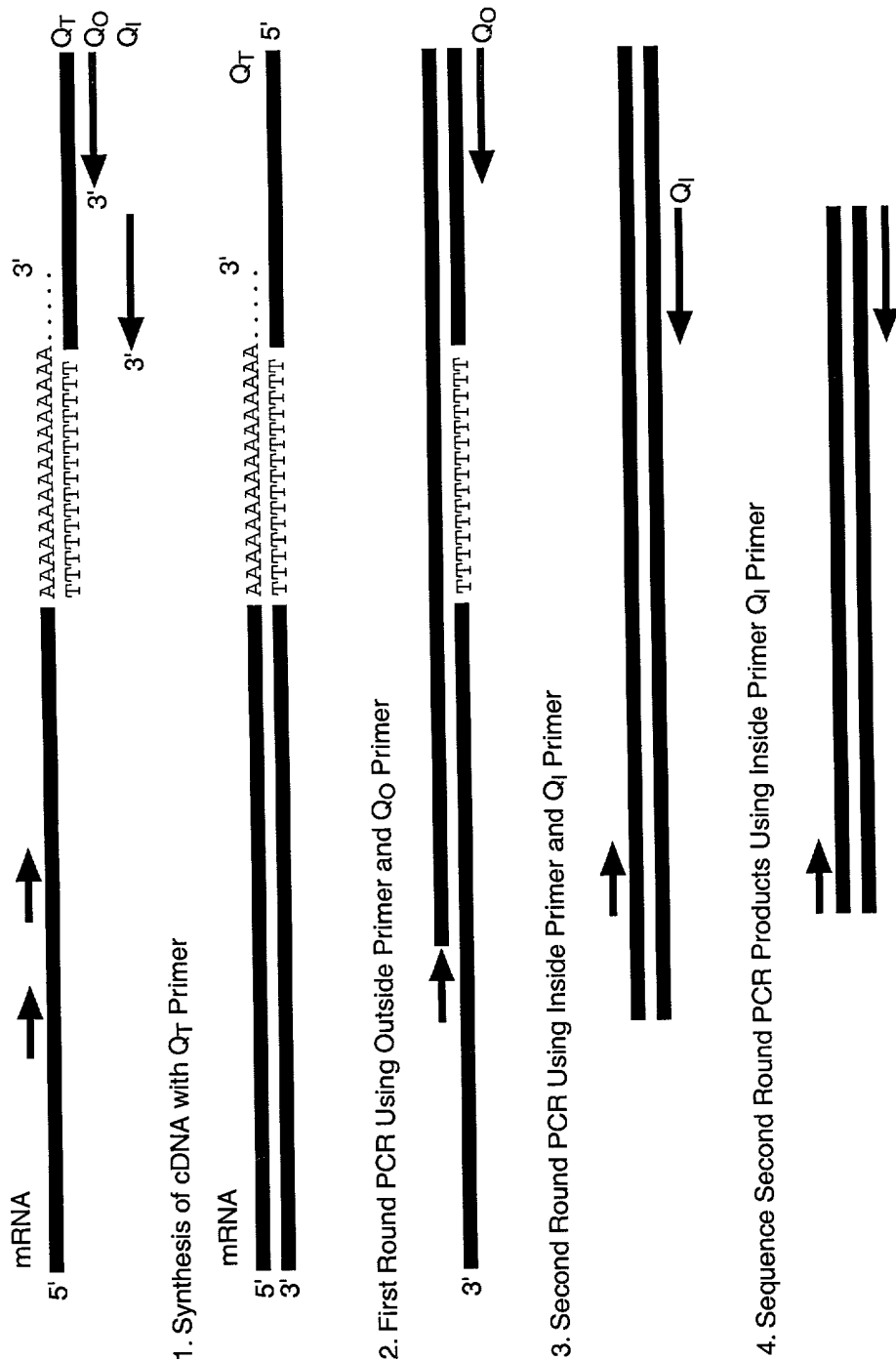


FIG. 37

+



+

- A. Genomic Libraries**
- Size Selected Libraries from P. Nurese
    - 3~4 kb
    - 5~7 kb
    - 7~8 kb
    - 11~12 kb
  - Libraries from J.A. Wise
    - Sau 3a Partial Digest
    - Hind III Partial Digest
- cDNA Libraries**
- GAD (Gal Activation Domain) Library
  - REP Library from R. Allshire
  - REP81ES Library (old)
  - REP81ES Library (new)
  - REP41ES Library

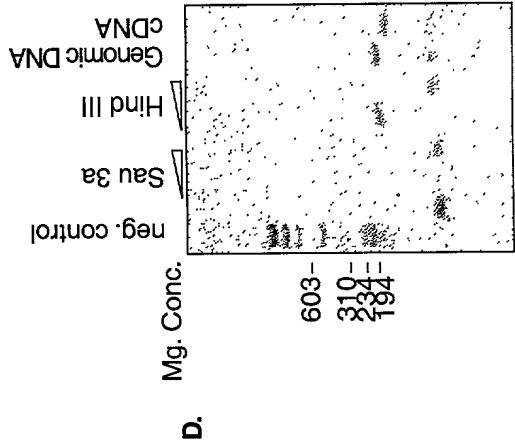
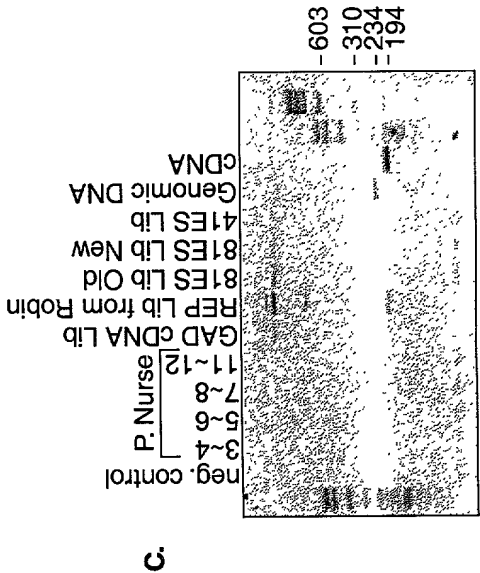
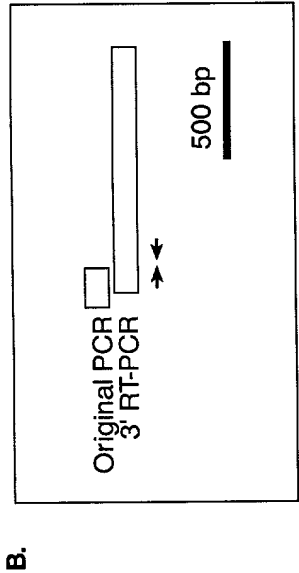


FIG. 38

+

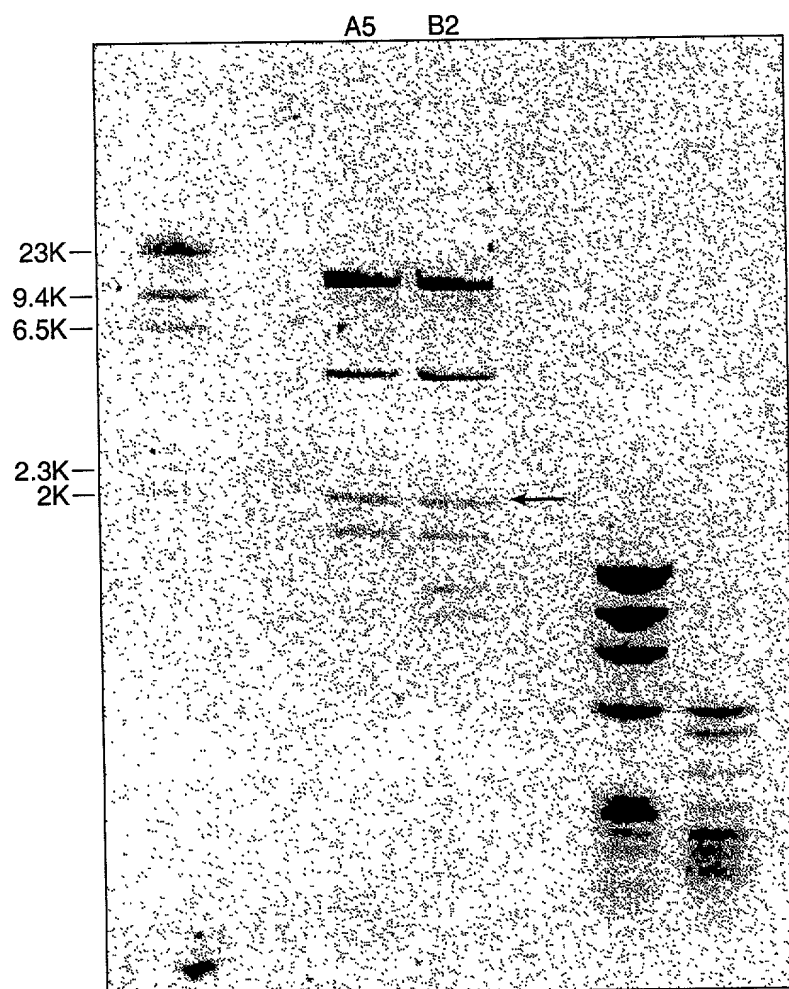


FIG. 39

09766253-082001

+

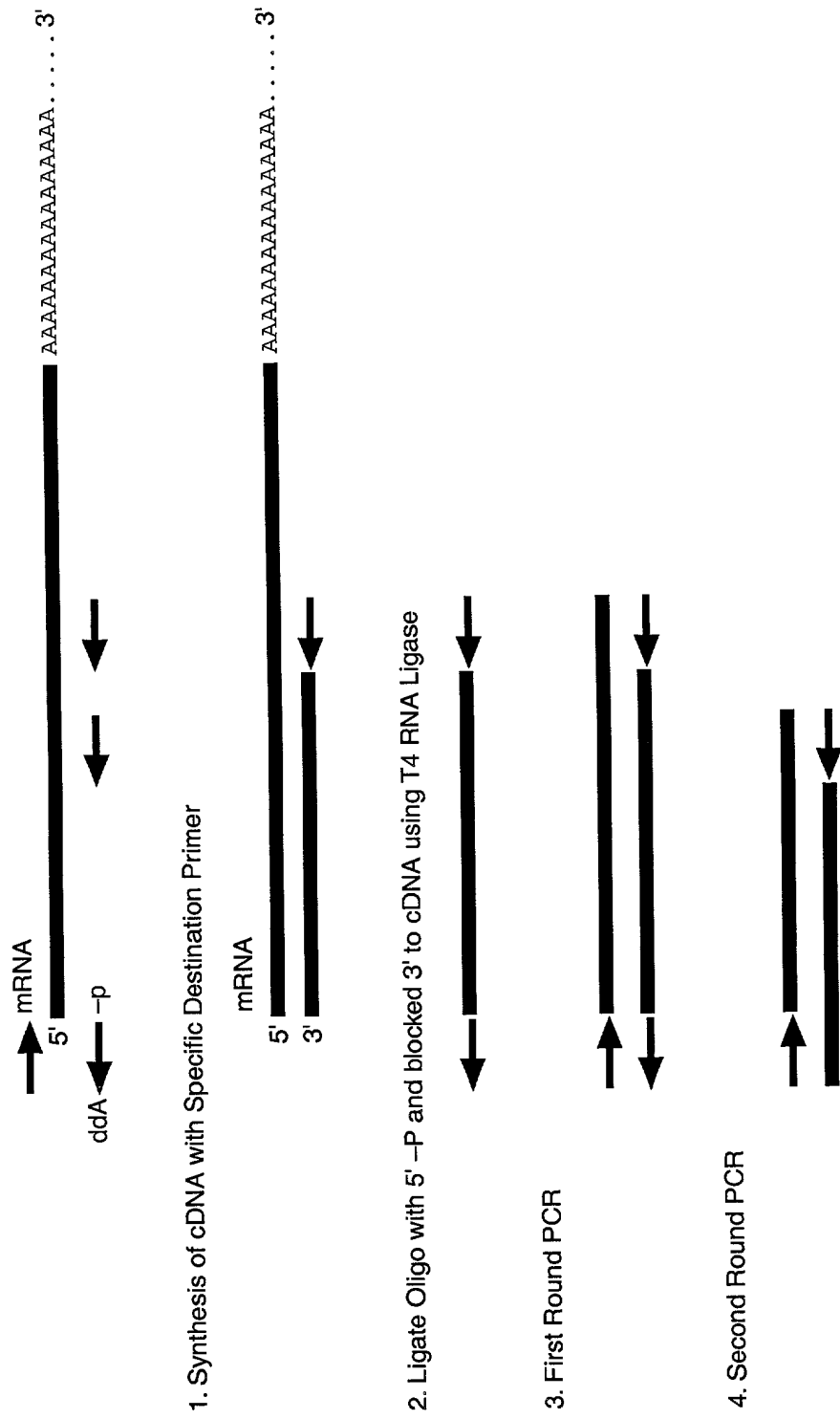


FIG. 40

+

+

				Motif O	
S.p. Tez1p	(429) .	WLYNSFIIPILQSFYIT	ESSDLNRRTVYFRKDIW	...	(35)...
S.c. Est2p	(366) .	WLFRQLIPKIIQTFYCTE	ISSTVT-IVYFRHDTW	...	(35)...
E.a. p123	(441) .	WIFEDLVVSLIRCFYVTE	QQKSYKTYIRKNIW	...	(35)...
		*** **	*	*	*
		Motif 1	Motif 2	K	
		p hh h K	hr h	R	
S.p. Tez1p		AVIRLLPKK--NTFRLITN	-LRKRF	...	(61)...
S.c. Est2p		SKMRIIPKKSNNEFRIIA	IPCRGAD	...	(62)...
E.a. p123		GKLRLLPKK--TTFRPIMT	FNKKIV	...	(61)...
		* ***	** *		
		Motif 3(A) AF			
		h hDh GY	h		
S.p. Tez1p		KKYFVRIDIKSCYDRIKQ	DLMFRIK	...	(89)...
S.c. Est2p		ELYFMKFDVKSCYDSIPR	MECMRIK	...	(75)...
E.a. p123		KLFFATMDIEKCYDSVNRE	KLSTFLK	...	(107)...
		* *	***	*	
		Motif 4(B')			
		hpQG	pp hh	h	
S.p. Tez1p		YLQKVGPQGSSILSSFLCH	FYMEDLIDEYLSF	...	(6)...
S.c. Est2p		YIREDGLFQGSSLSAPIVD	LVYDDLLEFYSEF	...	(8)...
E.a. p123		YKQTKGIPQGLCVSSILSS	FYATLEESSLGF	...	(14)...
		* *	** *	*	
		Y Motif 5(C)			
		h F DDhhh			
S.p. Tez1p		VLLRVVDDFLFITVNNKDA	KKFLNLSLRGFEKHNF	STLEKTVINFENS	.(205)
S.c. Est2p		LILKLADDFLIISTDQQQV	INIKKIAMGGFQKYN	AKANRDKILAVSSQS	.(173)
E.a. p123		LLMRLTDDYLLITTTQENNA	VLFIIEKLINVSRENG	FKFNMKKLQTSFFPLS	.(209)
		** *	*	*	*
		Motif 6(D)			
		Gh h	ck	h	

FIG. 41

+

FIG. 42

FIG. 42

Sp_Tip1p	219	W	N	S	I	S	I	S	R	F	S	I	F	Y	R	S	S	Y	K	F	K	Q	D	L	Y	F	N	L	H	S	I	C	D	251	
Sc_Est2p	184	N	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	K	Q	F	L	H	K	L	N	I	N	S	S	F	F	P	200	
Ea_p123	218	N	E	K	-	-	D	H	F	L	N	I	N	V	P	N	W	N	M	K	S	R	T	R	I	F	Y	C	T	H	F	N	248		
Sp_Tip1p	252	R	N	T	V	H	M	W	L	Q	W	I	F	P	R	Q	F	G	L	I	N	A	F	Q	V	K	Q	L	H	K	V	I	P	L	284
Sc_Est2p	201	-	-	-	-	-	-	-	-	Y	S	K	I	L	P	S	S	-	-	-	S	I	K	K	L	T	D	L	R	E	A	I	F	P	223
Ea_p123	249	R	-	-	-	-	-	-	-	N	N	Q	F	F	K	K	H	E	F	V	S	N	K	N	I	S	A	M	D	R	A	Q	T	I	275
Sp_Tip1p	285	V	S	-	-	-	Q	S	T	V	V	P	K	R	L	L	K	V	Y	P	L	I	E	Q	T	A	K	R	L	H	R	I	S	313	
Sc_Est2p	224	T	N	-	-	-	L	V	K	I	P	Q	R	L	K	V	R	I	N	L	T	L	Q	K	L	L	K	R	H	K	R	L	N	252	
Ea_p123	276	F	T	N	I	F	R	F	N	R	I	R	K	L	K	D	K	V	I	E	K	I	A	Y	M	L	E	K	V	K	D	F	N	308	
Sp_Tip1p	314	L	S	K	V	Y	N	H	Y	C	P	Y	I	D	-	T	H	D	D	E	K	I	L	S	Y	S	L	K	P	N	Q	-	-	342	
Sc_Est2p	253	Y	V	S	I	L	N	S	I	C	P	P	L	E	G	T	V	L	D	L	S	H	L	S	R	Q	S	P	K	E	R	-	-	282	
Ea_p123	309	F	N	Y	Y	L	T	K	S	C	P	L	P	E	N	W	R	E	R	K	Q	I	E	N	L	I	N	K	T	R	E	E	K	341	
Sp_Tip1p	343	-	-	-	-	-	-	-	-	-	-	-	-	-	-	V	F	A	F	L	R	S	I	L	V	R	V	F	P	K	L	I	359		
Sc_Est2p	283	-	-	-	-	-	-	-	-	-	-	-	-	-	-	V	L	K	F	I	I	V	I	L	Q	K	L	P	Q	E	M	-	299		
Ea_p123	342	S	K	Y	Y	E	E	L	F	S	Y	T	T	D	N	K	C	V	T	Q	F	I	N	E	F	F	Y	N	I	L	P	K	D	F	374
Sp_Tip1p	360	W	G	N	Q	R	I	F	E	I	I	L	K	D	L	E	T	F	L	K	L	S	R	Y	E	S	F	S	L	H	Y	L	M	S	392
Sc_Est2p	300	F	G	S	K	K	N	K	G	K	I	I	K	N	L	N	L	L	L	S	L	P	L	N	G	Y	L	P	F	D	S	L	L	K	332
Ea_p123	375	L	T	G	-	R	N	R	K	N	F	Q	K	K	V	K	Y	V	E	L	N	K	H	E	L	I	H	K	N	L	L	L	E	-	406
Sp_Tip1p	393	N	I	K	I	S	E	I	E	W	L	V	L	G	K	R	S	N	A	K	M	C	L	S	D	F	E	K	R	K	Q	I	F	A	425
Sc_Est2p	333	K	L	R	L	K	D	F	R	W	L	F	I	S	-	-	D	I	W	F	T	K	H	N	F	E	N	L	N	Q	L	A	I	362	
Ea_p123	407	K	I	N	T	R	E	I	S	M	M	Q	V	E	T	S	-	A	K	H	F	Y	Y	F	D	H	E	N	-	I	Y				

FIG. 42  
(CONTINUED)

+

A.	Sp_Tip1p	426	E	F	I	Y	W	L	Y	N	S	F	I	I	P	I	L	Q	S	F	F	Y	I	T	E	S	S	D	L	R	N	R	T	V	Y	458
	Sc_Est2p	363	C	F	I	S	W	L	F	R	Q	L	I	P	K	I	I	Q	T	F	F	Y	C	T	E	I	S	S	T	V	T	-	I	V	Y	394
	Ea_p123	438	K	L	L	R	W	I	F	E	D	L	V	V	S	L	I	R	C	F	F	Y	V	T	E	Q	Q	K	S	Y	S	K	T	Y	Y	470
	Sp_Tip1p	459	F	R	K	D	I	W	K	L	L	C	R	P	F	I	T	S	M	K	M	E	A	F	E	K	I	N	E	N	N	V	R	M	D	491
	Sc_Est2p	395	F	R	H	D	T	W	N	K	L	I	T	P	F	I	V	E	Y	E	K	T	Y	L	V	E	N	N	V	C	R	N	H	N	S	427
	Ea_p123	471	Y	R	K	N	I	W	D	V	I	M	K	M	S	I	A	D	L	K	K	E	T	L	A	E	V	Q	E	K	E	V	E	E	W	503
	Sp_Tip1p	492	T	Q	K	T	T	L	P	A	V	I	R	L	L	P	K	K	-	-	N	T	F	R	L	I	T	N	L	R	K	R	F	L	522	
	Sc_Est2p	428	Y	T	L	S	N	F	N	H	S	K	M	R	I	I	P	K	K	S	N	E	F	R	I	I	A	I	P	C	R	G	A	D	460	
	Ea_p123	504	K	K	S	L	G	F	A	P	G	K	L	R	L	I	P	K	K	-	-	T	T	F	R	P	I	M	T	F	N	K	I	V	534	
	Sp_Tip1p	523	I	K	M	G	S	N	K	K	M	L	V	S	T	N	Q	T	L	R	P	V	A	S	I	L	K	H	L	I	N	E	-	-	552	
	Sc_Est2p	461	E	E	E	-	-	F	T	I	Y	K	E	N	H	K	N	A	I	Q	P	T	Q	K	I	L	E	Y	L	R	N	K	R	P	T	491
	Ea_p123	535	N	S	D	-	-	R	K	T	T	K	L	T	T	N	T	K	L	N	S	H	L	M	L	K	T	L	K	N	R	-	M	F	564	
	Sp_Tip1p	553	E	S	S	G	I	P	F	N	L	E	V	Y	M	K	L	L	T	F	K	K	D	L	L	K	H	R	M	F	G	R	-	K	K	584
	Sc_Est2p	492	S	F	T	K	I	Y	S	P	T	Q	I	A	D	R	I	K	E	F	K	Q	R	L	L	K	K	F	N	N	V	L	P	E	L	524
	Ea_p123	565	K	D	P	F	G	F	A	V	F	N	Y	D	D	V	M	K	K	Y	E	E	F	V	C	K	W	K	Q	V	G	P	K	L	597	
	Sp_Tip1p	585	Y	F	V	R	I	D	I	K	S	C	Y	D	R	I	K	Q	D	L	M	F	R	I	V	K	K	L	K	D	P	E	-	F	616	
	Sc_Est2p	525	Y	F	M	K	F	D	V	K	S	C	Y	D	S	I	P	R	M	E	C	M	R	I	L	K	D	A	L	K	N	E	N	G	F	557
	Ea_p123	598	F	F	A	T	M	D	I	E	K	C	Y	D	S	V	N	R	E	K	L	S	T	F	L	K	T	T	K	L	L	S	S	D	F	630
	Sp_Tip1p	617	V	I	R	K	Y	A	T	I	H	A	T	S	D	R	A	T	K	N	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	634
	Sc_Est2p	558	F	V	R	S	Q	Y	F	F	N	T	N	T	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	570	
	Ea_p123	631	W	I	I	M	T	A	Q	I	L	K	R	K	N	I	V	I	D	S	K	N	F	R	K	E	M	K	D	Y	F	R	Q	K	663	

FIG. 42  
(CONTINUED)

+

+

A.	Sp_Tip1p	635	F	V	S	E	A	F	S	Y	F	D	M	V	P	F	E	K	V	V	Q	L	S	-	-	M	K	T	S	D	T	L	F	V	665					
	Sc_Est2p	571	-	-	-	-	-	-	-	-	-	-	-	V	L	K	L	F	N	V	V	N	A	S	R	-	-	V	P	K	P	Y	E	I	591					
	Ea_p123	664	F	Q	K	I	A	L	E	G	Q	Y	P	T	L	F	S	V	L	E	N	E	Q	N	D	L	N	A	K	K	T	L	I	V	696					
	Sp_Tip1p	666	D	F	V	D	Y	W	T	K	S	S	S	E	I	F	K	M	L	K	E	H	L	S	G	H	I	V	K	I	G	N	S	Q	Y	698				
	Sc_Est2p	592	D	N	V	R	T	V	H	L	S	N	Q	D	V	I	N	V	V	E	M	E	I	F	K	T	A	L	W	V	E	D	K	C	Y	624				
	Ea_p123	697	E	A	K	Q	R	N	Y	F	K	K	D	N	L	L	Q	P	V	I	N	I	C	Q	Y	N	Y	I	N	F	N	G	K	F	Y	729				
	Sp_Tip1p	699	L	Q	K	V	G	I	P	Q	G	S	I	L	S	S	F	L	C	H	F	Y	M	E	D	L	I	D	E	Y	L	S	F	T	K	731				
	Sc_Est2p	625	I	R	E	D	G	L	F	Q	G	S	S	L	S	A	P	I	V	D	L	V	Y	D	D	L	L	E	F	Y	S	E	F	K	A	657				
	Ea_p123	730	K	Q	T	K	G	I	P	Q	G	L	C	V	S	S	I	L	S	S	F	Y	A	T	L	E	E	S	S	L	G	F	L	R	762					
	Sp_Tip1p	732	K	K	G	-	-	-	-	-	-	-	-	-	-	-	-	-	S	V	L	L	R	V	V	D	D	F	L	F	I	T	V	N	K	K	D	A	K	756
	Sc_Est2p	658	S	P	S	Q	D	-	-	-	-	-	-	-	-	-	-	-	T	L	I	L	K	L	A	D	D	F	L	I	I	S	T	D	Q	Q	V	I	N	684
	Ea_p123	763	D	E	S	M	N	P	E	N	P	N	V	N	L	L	M	R	L	T	D	D	Y	L	L	I	T	T	Q	E	N	N	A	V	L	795				
	Sp_Tip1p	757	F	L	N	L	S	L	R	G	F	E	K	H	N	F	S	T	S	L	E	K	T	V	I	N	F	E	N	S	N	G	-	-	-	-	-	786		
	Sc_Est2p	685	I	K	K	L	A	M	G	G	F	Q	K	Y	N	A	K	A	N	R	D	K	I	L	A	V	S	S	Q	S	D	-	-	-	-	-	-	713		
	Ea_p123	796	F	I	E	K	L	I	N	V	S	R	E	N	G	F	K	F	N	M	K	L	Q	T	S	E	F	P	L	S	P	S	K	F	A	828				
	Sp_Tip1p	787	-	-	-	I	I	N	N	T	F	F	N	E	S	K	K	R	M	P	F	F	G	F	S	V	N	M	R	S	L	D	T	L	L	816				
	Sc_Est2p	714	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	M	H	I	F	V	K	E	L	E	V	W	K	H	S	S	T	M	739				
	Ea_p123	829	K	Y	G	M	D	S	V	E	E	Q	N	I	V	Q	D	Y	C	D	W	I	G	I	S	I	D	M	K	T	L	A	L	M	P	861				
	Sp_Tip1p	817	A	C	P	K	I	D	E	A	L	F	N	S	T	S	V	E	L	T	K	H	M	G	K	S	F	F	Y	K	I	L	R	S	S	849				
	Sc_Est2p	740	N	N	F	H	I	R	S	K	S	S	K	G	I	F	R	S	L	I	A	L	F	N	T	R	I	S	Y	K	T	I	D	T	N	772				
	Ea_p123	862	N	I	N	L	R	I	E	G	I	L	C	T	L	N	L	N	M	Q	T	K	K	A	S	M	W	L	K	K	L	K	S	F	894					

FIG. 42  
(CONTINUED)

+



+

A.																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
----	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

FIG. 42  
(CONTINUED)

+

+

B.		
Sp_Tip1p	1	- - - - - MTEHHTPKSRILRFLLENQYVYLCT 24
Sc_Est2p	1	- - - - - - - - - - - - - - - MKILFEF 7
Ea_p123	1	MEVDVDNQADNHGISHALKTCEEIKEAKTLYSW 33
Sp_Tip1p	25	LNDYVQLVLRGSPA[S]Y[SN]ICERLRSRVQTSFS 57
Sc_Est2p	8	IQDKLDIDLQTN--S[TYK]--ENLKC[GHF]NGLD 35
Ea_p123	34	IQKVIRCRNQSQ--S[HYK]--DLEDIKIFAQTN 61
Sp_Tip1p	58	IFLHSTVVGFDSKPD[EGV]QFSSPKCSQSELI[AN] 90
Sc_Est2p	36	EILTTTCFALPNSR-KIALPCLPGDLSHKAVI[DH] 67
Ea_p123	62	I VATPRDYNEEDFKVIARKEVFSTGLMIE[LI]DK 94
Sp_Tip1p	91	VVKQMFDESFERRR-NLLMK[G]FSMNHEDFRAMH 122
Sc_Est2p	68	CIIYLLTGELYN--NVLTFG[YK]IARNED-- 93
Ea_p123	95	CLVELLSSSDVSDRQKLQC[F]G[FQ]LKGNQ-- 122
Sp_Tip1p	123	VNGVQNDLVSTFPN[YL]ISILESKN[WQ]LLLEI[G] 155
Sc_Est2p	94	- - - VNNSLFC[HS]ANVNVTTLLKGA[AWK]MFHSLV[G] 123
Ea_p123	123	- - - LAKTHLLTALSTQKQYFFQDEWNQVRAMI[G] 152
Sp_Tip1p	156	SDAMHY[L]LSKGSIFEALPNDNYLQ[ISG]IPLFKN 188
Sc_Est2p	124	TYAFVDLLIN[YT]VIQFN-GQFFTTQ[IVGN]RCNEP 155
Ea_p123	153	NELFRHL[YTK]YLI[FQRT]SEGTLVQ[FCG]NNVFDH 185
Sp_Tip1p	189	NVFEETVSKKRKRRTIETSI[TON]--KSARKE[VS] 218
Sc_Est2p	156	HLPPKWVQ--RSSSSSATAAQI--KQLTEPVT 183
Ea_p123	186	LKVNDKFDK-KQKGGAA[DMNE]PRCCSTCKYNV[K] 217

FIG. 42  
(CONTINUED)

+

+

<b>B.</b>			
<b>Sp_Tip1p</b>	219	WNSISISRFSIF YRSSYKKFKQDLYFNLSHSICD	251
<b>Sc_Est2p</b>	184	N - - - - -	200
<b>Ea_p123</b>	218	NEK - - DHFLNNINVPNWNMKSRTIRIFYCTHFN	248
<b>Sp_Tip1p</b>	252	RNTVHMWLQWIFPRQFGLINAFQVKQLHKV IPL	284
<b>Sc_Est2p</b>	201	- - - - - YSKILPSS - - - SIKKLTDLREAIFP	223
<b>Ea_p123</b>	249	R - - - - - NNQFFKKHEFVSNNKI SAMDRAQTI	275
<b>Sp_Tip1p</b>	285	VS - - - - QSTVVPKRL LKVYPLIEQTAKRRLHRI S	313
<b>Sc_Est2p</b>	224	TN - - - - - LVKIPQRLKVRINLTQLKLLKRHKRLN	252
<b>Ea_p123</b>	276	FTNIFRFNRI RKKLKDKVIEKIA YMLEKVKDFN	308
<b>Sp_Tip1p</b>	314	LSKVYNHYCPYID - THDDEKILSYSLKPNQ - - -	342
<b>Sc_Est2p</b>	253	YVSI LNSICPPLEGTVLDLSHLSRQSPKER - - -	282
<b>Ea_p123</b>	309	FNYYLT KSCPLPENWRERKQKIENL INKTREEK	341
<b>Sp_Tip1p</b>	343	- - - - - - - - - - VFALRSILVRVFPKLI	359
<b>Sc_Est2p</b>	283	- - - - - - - - - - VLKFIIVILQKLLPQEM	299
<b>Ea_p123</b>	342	SKYYEELFSYTTDNKCVTQF INEFFYNI LPKDF	374
<b>Sp_Tip1p</b>	360	WGNQRIFEIILKDL ETFLKL SRYESFSLHYLMS	392
<b>Sc_Est2p</b>	300	FGSKKNKGKIIKNLNL LSLPLNGYLPFD SLLK	332
<b>Ea_p123</b>	375	LTG - RNRKNFQKKVKKYVELNKH ELIHKNNLLE	406
<b>Sp_Tip1p</b>	393	NIKISEIEWLVLGKRSNAKMCLSDFFEKRKQIFA	425
<b>Sc_Est2p</b>	333	KLRLKDFR WLFIS - - - DIWFTKHN FENLNQLAI	362
<b>Ea_p123</b>	407	KINTREISWMQVETS - AKHFYYFDHE - IYVLW	437

FIG. 42  
(CONTINUED)

+

B.			
Sp_Tip1p	426	EF I Y W L Y N S F I I P I L Q S F F Y I T E S S D L R N R T V Y	458
Sc_Est2p	363	CF I S W L F R Q L I P K I I Q T F F Y C T E I S S T V T - I V Y	394
Ea_p123	438	K L L R W I F E D L V V S L I R C F F Y V T E Q Q K S Y S K T Y Y	470
Sp_Tip1p	459	F R K D I W K L L C R P F I T S M K M E A F E K I N E N N V R M D	491
Sc_Est2p	395	F R H D T W N K L I T P F I V E Y F K T Y L V E N N V C R N H N S	427
Ea_p123	471	Y R K N I W D V I M K M S I A D L K K E T L A E V Q E K E V E E W	503
Sp_Tip1p	492	T Q K T T L P P A V I R L L P K K - - N T F R L I T N L R K R F L	522
Sc_Est2p	428	Y T L S N F N H S K M R I I P K K S N N E F R I I A I P C R G A D	460
Ea_p123	504	K K S L G F A P G K L R L I P K K - - T T F R P I M T F N K K I V	534
Sp_Tip1p	523	I K M G S N K K M L V S T N Q T L R P V A S I L K H L I N E - - -	552
Sc_Est2p	461	E E E - - F T I Y K E N H K N A I Q P T Q K I L E Y L R N K R P T	491
Ea_p123	535	N S D - - R K T T K L T T N T K L L N S H L M L K T L K N R - M F	564
Sp_Tip1p	553	E S S G I P F N L E V Y M K L L T F K K D L L K H R M F G R - K K	584
Sc_Est2p	492	S F T K I Y S P T Q I A D R I K E F K Q R L L K K F N N V L P E L	524
Ea_p123	565	K D P F G F A V F N Y D D V M K K Y E E F V C K W K V G Q P K L	597
Sp_Tip1p	585	Y F V R I D I K S C Y D R I K Q D L M F R I V K K K L K D P E - F	616
Sc_Est2p	525	Y F M K F D V K S C Y D S I P R M E C M R I L K D A L K N E N G F	557
Ea_p123	598	F F A T M D I E K C Y D S V N R E K L S T F L K T T K L L S S D F	630
Sp_Tip1p	617	V I R K Y A T I H A T S D R A T K N - - - - - - - - - - - - -	634
Sc_Est2p	558	F V R S Q Y F F N T N T G - - - - - - - - - - - - - - - -	570
Ea_p123	631	W I M T A Q I L K R K N N I V I D S K N F R K K E M K D Y F R Q K	663

FIG. 42  
(CONTINUED)

+

B.		
Sp_Tip1p	635	FVSEAFSYFDMVPFEK[V]VQLLS--MKTSDT[L]FV 665
Sc_Est2p	571	-----VLKLFN[V]VNASR--VPKPYEL[Y]I 591
Ea_p123	664	FQKIALEGGQYPTLFS[V]LENEQNDLNAKKT[L]IV 696
Sp_Tip1p	666	DFVDYWTKSSSEIFKMLLKEHLSGHI VKIGNSQ[Y] 698
Sc_Est2p	592	DNVRTVHL SNQDVINVVEMEIFKTA LWVEDKCY 624
Ea_p123	697	EAKQRNYFKKDNLLQPVINICQYNYINFNKGK[Y] 729
Sp_Tip1p	699	LQKV[G]IPQG[S]ILSSFLCHFYMED[L]IDEYLS[FT]K 731
Sc_Est2p	625	IREDGLFQGS[SL]SAPIVDLVYDD[L]LEFYSE[FF]KA 657
Ea_p123	730	KQTKG[IP]QG[LCV]SSILSSFYAT[L]EES[SL]FLR 762
Sp_Tip1p	732	KKG-----SVLLRVVDD[F]L[F]ITVNKKDAKK 756
Sc_Est2p	658	SPSQD-----TLILKLADD[F]L[I]ISTDQQQVIN 684
Ea_p123	763	DESMNPENPNVNLMLRLTDD[Y]L[L]ITTQENNAVL 795
Sp_Tip1p	757	FLNLSLRGFEKHNFSSTLE[K]TVINFEN[SN]G--- 786
Sc_Est2p	685	IKKLAMGGFQKYNAKANRD[K]ILAVSSQ[S]D--- 713
Ea_p123	796	FIEKLINVSRENGFKFNMK[K]LQTSFPL[S]PSKFA 828
Sp_Tip1p	787	---IINNNTFFNESKKRMPFFGFSVNMRSLDTLL 816
Sc_Est2p	714	---DDTTVIQFCA---MHIFVKELEVWKHSS[ST]M 739
Ea_p123	829	KYGMDSVEEQNI VQDYCDWIGISIDMKTLALMP 861
Sp_Tip1p	817	ACPKIDEALFNSTSVELTKHMGKSFFY[K]ILRSS 849
Sc_Est2p	740	NNFHIRSKSSKGI FRSLIALFNTRISYK[T]IDTN 772
Ea_p123	862	NINLRIEGLCTLNLMQTKKASMWLKK[K]KLSF 894

FIG. 42  
(CONTINUED)

+

+

B.

Sp_Tip1p	850	L	A	S	F	A	Q	V	F	I	D	I	T	H	N	S	K	F	N	S	C	C	N	I	Y	R	L	G	Y	S	M	C	M	R	882	
Sc_Est2p	773	L	N	S	T	N	T	V	L	M	Q	I	D	H	V	V	K	N	I	S	E	C	-	-	-	-	-	-	-	-	-	-	-	-	-	793
Ea_p123	895	L	M	N	N	I	T	H	Y	F	R	K	T	I	T	T	E	D	F	A	N	K	T	L	N	K	L	F	I	S	G	G	Y	K	927	
Sp_Tip1p	883	A	Q	A	Y	L	K	R	M	K	D	I	F	I	P	Q	R	M	F	I	T	D	L	L	N	V	I	G	R	K	I	W	K	K	915	
Sc_Est2p	794	-	-	-	Y	K	S	A	F	K	D	L	S	I	N	-	-	V	T	Q	N	M	Q	F	H	S	F	L	Q	R	I	I	E	M	821	
Ea_p123	928	Y	M	Q	C	A	K	E	Y	K	D	H	F	K	K	N	L	A	M	S	S	M	I	D	L	E	V	S	K	I	I	Y	S	V	960	
Sp_Tip1p	916	L	A	E	I	L	G	Y	T	S	R	R	F	L	S	S	A	E	V	K	W	L	F	C	L	G	M	R	D	G	L	K	P	S	948	
Sc_Est2p	822	T	V	S	G	C	P	I	T	K	C	D	P	L	I	E	Y	E	V	R	F	T	I	L	N	G	F	L	E	S	L	S	S	N	854	
Ea_p123	961	T	R	A	F	F	K	Y	L	V	C	N	I	K	D	T	I	F	G	E	E	H	Y	P	D	F	F	L	S	T	L	K	H	F	993	
Sp_Tip1p	949	F	K	Y	H	P	C	F	E	Q	L	I	Y	Q	F	Q	S	L	T	D	L	I	K	P	L	R	P	V	L	R	Q	V	L	F	981	
Sc_Est2p	855	T	S	-	-	-	-	-	-	-	-	-	-	-	K	F	K	D	N	I	I	L	R	K	E	I	Q	H	L	Q	A	Y	I	Y	877	
Ea_p123	994	I	E	I	F	S	-	-	-	-	-	-	-	-	T	K	K	Y	I	F	N	R	V	C	M	I	L	K	A	K	E	A	K	L	1023	
Sp_Tip1p	982	L	H	R	R	I	A	D	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	988		
Sc_Est2p	878	I	Y	I	H	I	V	N	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	884		
Ea_p123	1024	Q	S	L	I	Q	Y	D	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1031		

FIG. 42  
(CONTINUED)

+

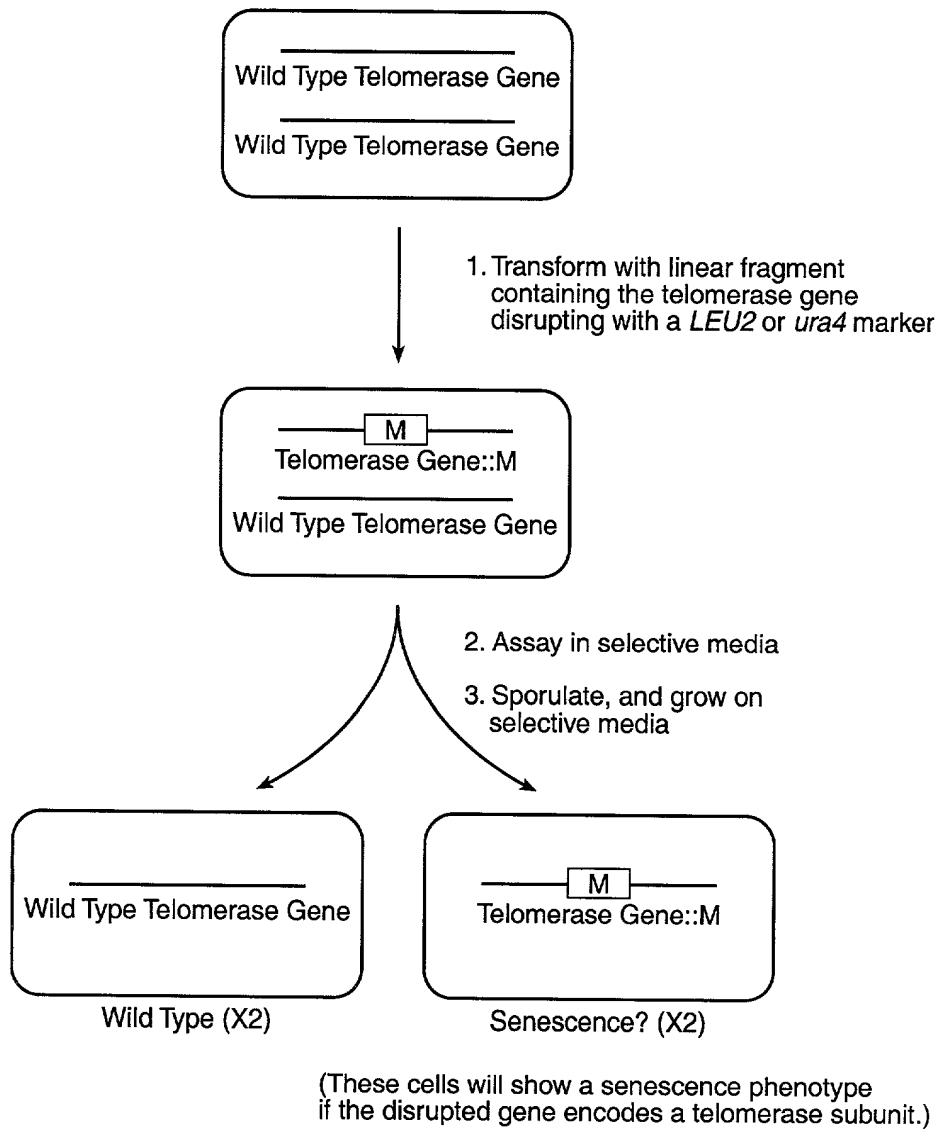


FIG. 43

+

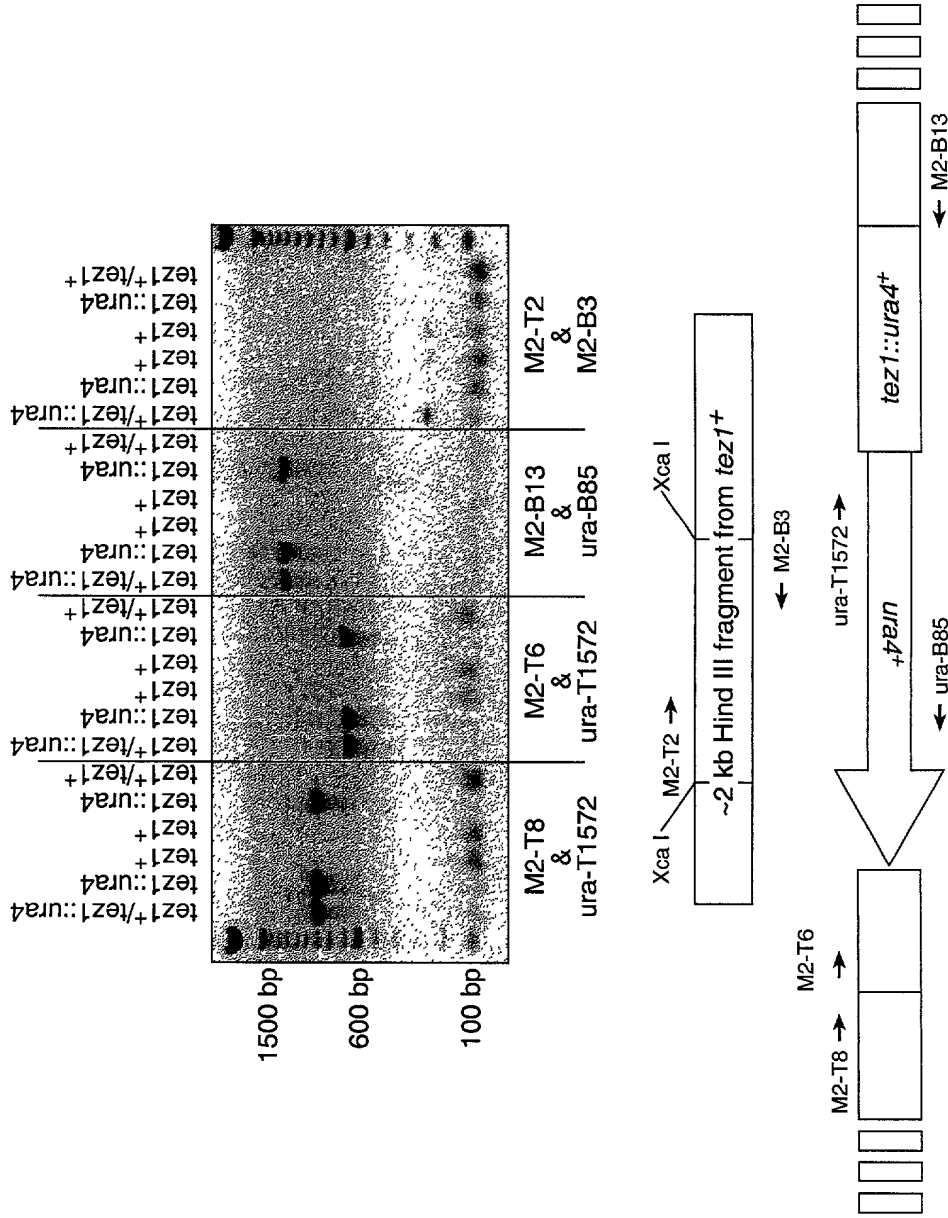


FIG. 44

+



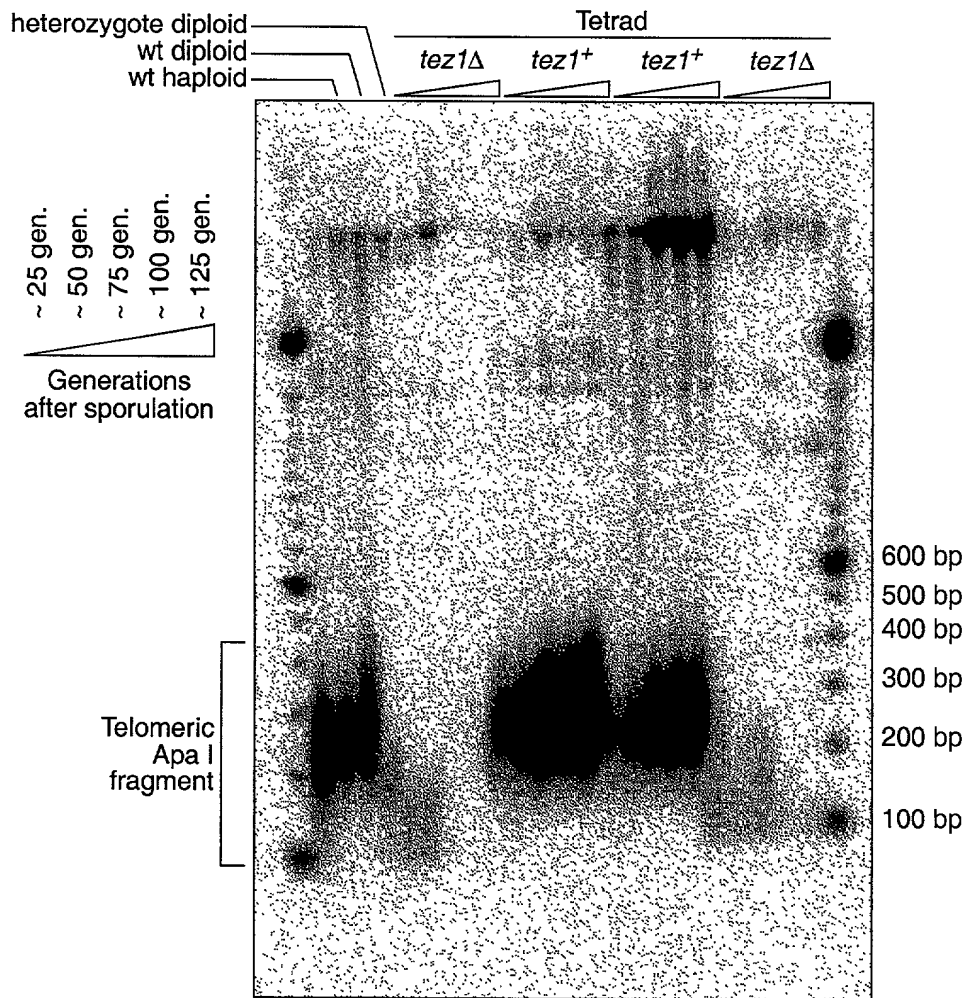


FIG. 45

+

1 ggtaccgatttactttcttcttcaataagtaattgcttctcgaacgctcctaaatctctggaataatttttacaaga 80  
 81 actcaataacaataccaagtcaaatccaattatgaaggtgtatttagtgatcgataaatattttctattttatcggtcgta 160  
 161 ccaagtataaggacaaaagaacaactctccctcccaagacttttactatttatttaatttcttcaaatatattcgc 240  
 241 ggttcgttacttttaatacgttggtactgttttagctgacttcttagccaaacggtgtttctaccccgctcatcgatat 320  
 321 agctctggagtagctcacagaaactctacaatctctcgaactatgagactatattagattctacagtcctgcatattc 400  
 401 ttaacatggagccttacactttagatgagtcacgctgcgcatgaggatatttgggtatcatccaacggtttgccttgaaga 480  
 481 gttgataattatttgcaaatcatgctccttagtgggtgaatccgcgaagtttttttgatgctgcacacgctctagcatg 560  
 561 attgagatatttcaaaaatttctatccactacaactcctttaacgcggtttttattttctattttctattctcatgtgtt 640  
 641 ccaaatatgtatcatcctcgtattaggctttttccgttttactcctggaatcgtaaccttttctactattccccctaatga 720  
 721 ataactaaaattagtttcgcttataattgatagtagtagaagatttggtgactctactcgtgtaattgtattagtttaaa 800  
 801 gatactttgcaaacatttattagctatcattataaaaaaacctcattataaaataatcaatcaatatttgcggtc 880  
 881 actatttttaaaacgcttatgatcagtaggacactttgcataatatatgattatgcttaattggttacttgaacttgc 958  
  
 959 ATG ACC GAA CAC CAT ACC CCC AAA AGC AGG ATT CTT CGC TTT CTA GAG AAT CAA TAT GTA 1018  
 1 M T E H T P K S R I L R F L E N Q Y V 20  
  
 1019 TAC CTA TGT ACC TTA AAT GAT TAT GTA CAA CTT GTT TTG AGA GGG TCG CCG GCA AGC TCG 1078  
 21 Y L C T L N D Y V Q L V L R G S P A S S 40  
  
 1079 TAT AGC AAT ATA TGC GAA CGC TTG AGA AGC GAT GTA CAA ACG TCC TTT TCT ATT TTT CTT 1138  
 41 Y S N I C E R L R S D V Q T S F S I F L 60  
  
 1139 CAT TCG ACT GTA GTC GGC TTC GAC AGT AAG CCA GAT GAA GGT GTT CAA TTT TCT TCT CCA 1198  
 61 H S T V V G F D S K P D E G V Q F S S P 80  
  
 1199 AAA TGC TCA CAG TCA GAG gtatatatatatttttggtttttatttttctatttcgggatagctaataatatgggcag 1272  
 81 K C S Q S E 86  
  
 1273 CTA ATA GCG AAT GTT GTA AAA CAG ATG TTC GAT GAA AGT TTT GAG CGT CGA AGG AAT CTA 1332  
 87 L I A N V V K Q M F D E S F E R R N L 106  
  
 1333 CTG ATG AAA GGG TTT TCC ATG gtaagggtatttctaattgtgaaatatttaccctgcaattactgttttcaaagaga 1405  
 107 L M K G F S M 113  
  
 1406 ttgtattttaaccgataaag AAT CAT GAA GAT TTT CGA GCC ATG CAT GTA AAC GGA GTA CAA AAT 1469  
 114 N H E D F R A M H V N G V Q N 128

FIG. 46

+

$+$ 

FIG. 46  
(CONTINUED)

+

52/59

2268 TTT GAG ATA ATA TTA AAA G gtaattgtataaaattttattaccactaaacgattttaccag AC CTC GAA ACT 2336  
366 F E I I L K D L E T 375

2337 TTC TTG AAA TTA TCG AGA TAC GAG TCT TTT AGT TTA CAT TAT TTA ATG AGT AAC ATA AAG 2396  
376 F L K L S R Y E S F S L H Y L M S N I K 395

2397 gtaatatgccaaaattttttaccattaataacaatcag ATT TCA GAA ATT GAA TGG CTA GTC CTT GGA 2465  
396 I S E I E W L V L G 405

2466 AAA AGG TCA AAT GCG AAA ATG TGC TTA AGT GAT TTT GAG AAA CGC AAG CAA ATA TTT GCG 2525  
406 K R S N A K M C L S D F E K R K Q I F A 425

2526 GAA TTC ATC TAC TGG CTA TAC AAT TCG TTT ATA ATA CCT ATT TTA CAA TCT TTT TTT TAT 2585  
426 E F I Y W L Y N S F I I P I L Q S F Y 445

2586 ATC ACT GAA TCA AGT GAT TTA CGA AAT CGA ACT GTT TAT TTT AGA AAA GAT ATT TGG AAA 2645  
446 I T E S S D L R N R T V Y F R K D I W K 465

2646 CTC TTG TGC CGA CCC TTT ATT ACA TCA ATG AAA ATG GAA GCG TTT GAA AAA ATA AAC GAG 2705  
466 L L C R P F I T S M K M E A F E K I N E 485

2706 gtatttttaaagtattttttgcaaaaagctaattttcag AAC AAT GTT AGG ATG GAT ACT CAG AAA ACT 2775  
486 N N V R M D T Q K T 495

2776 ACT TTG CCT CCA GCA GTT ATT CGT CTA TTA CCT AAG AAG AAT ACC TTT CGT CTC ATT ACG 2835  
496 T L P P A V I R L L P K K N T F R L I T 515

2836 AAT TTA AGA AAA AGA TTC TTA ATA AAG gtatttaatttttggtcatcaatgtactttacttctaattatta 2906  
516 N L R K R F L I K 524

2907 ttagcag ATG GGT TCA AAC AAA AAA ATG TTA GTC AGT ACG AAC CAA ACT TTA CGA CCT GTG 2967  
525 M G S N K K M L V S T N Q T L R P V 542

2968 GCA TCG ATA CTG AAA CAT TTA ATC AAT GAA GAA AGT AGT GGT ATT CCA TTT AAC TTG GAG 3027  
543 A S I L K H L I N E S S G I P F N L E 562

3028 GTT TAC ATG AAG CTT ACT TTT AAG AAG GAT CTT CTT AAG CAC CGA ATG TTT GG gtaaat 3088  
563 V Y M K L L T F K K D L L K H R M F G 581

FIG. 46  
(CONTINUED)

+

+

3089 tatataatgcgcgattcctcattatttaatttgcag G CGT AAG AAG TAT TTT GTA CGG ATA GAT ATA 3155  
582 R K K Y F V R I D I 591

3156 AAA TCC TGT TAT GAT CGA ATA AAG CAA GAT TTG ATG TTT CGG ATT GTT AAA AAG AAA CTC 3215  
592 K S C Y D R I K Q D L M F R I V K K L 611

3216 AAG GAT CCC GAA TTT GTA ATT CGA AAG TAT GCA ACC ATA CAT GCA ACA AGT GAC CGA GCT 3275  
612 K D P E F V I R K Y A T I H A T S D R A 631

3276 ACA AAA AAC TTT GTT AGT GAG GCG TTT TCC TAT T gtaagtttatttttcatggaattttttaacaa 3343  
632 T K N F V S E A F S Y F 643

3344 attcttttttttag TT GAT ATG GTG CCT TTT GAA AAA GTC GTG CAG TTA CTT TCT ATG AAA ACA 3405  
644 D M V P F E K V V Q L L S M K T 659

3406 TCA GAT ACT TTG TTT GTT GAT TTT GTG GAT TAT TGG ACC AAA AGT TCT TCT GAA ATT TTT 3465  
660 S D T L F V D F V D Y W T K S S E I F 679

3466 AAA ATG CTC AAG GAA CAT CTC TCT GGA CAC AAT GTT AAG gtataccaattgtgaattgtaataaca 3532  
680 K M L K E H L S G H I V K 692

3533 ctaatgaaactag ATA GGA AAT TCT CAA TAC CTT CAA AAA GTT GGT ATC CCT CAG GGC TCA 3593  
693 I G N S Q Y L Q K V G I P Q G S 708

3594 ATT CTG TCA TCT TTT TTG TGT CAT TTC TAT ATG GAA GAT TTG ATT GAT GAA TAC CTA TCG 3653  
709 I L S S F L C H F Y M E D L I D E Y L S 728

3654 TTT ACG AAA AAG AAA GGA TCA GTG TTG TTA CGA GTA GTC GAC GAT TTC CTC TTT ATA ACA 3713  
729 F T K K G S V L L R V V D D F L F I T 748

3714 GTT AAT AAA AAG GAT GCA AAA AAA TTT TTG AAT TTA TCT TTA AGA G gtgagttgctgctattcc 3777  
749 V N K K D A K K F L N L S L R G 764

3778 taagttctaaccgttggaag GA TTT GAG AAA CAC AAT TTT TCT ACG AGC CTG GAG AAA ACA GTA 3840  
765 F E K H N F S T S L E K T V 778

3841 ATA AAC TTT GAA AAT AGT AAT GGG ATA ATA AAC AAT ACT TTT TTT AAT GAA AGC AAG AAA 3900  
779 I N F E N S N G I I N N T F F N E S K K 798

FIG. 46  
(CONTINUED)

+

FIG. 46  
(CONTINUED)

+

4666 gctgaccccccaagcaagcatactataggtttctagtaaaagtaaaataatctcgttattagttttgattgacttgtct 4745  
 4746 ttatccttataacttttaagaaagattgacagtggttgctgactactgcccacatgcccatataacgggagtggttaaaca 4825  
 4826 ttaaaagtaatacatgaggttaactcctttcattagaataaggaaagtggttttctataatgaataatgccgcacta 4905  
 4906 atgcaaaaagacgaagattatcttcaacaaggggattaaagcatatccgaaggaaagagagagataatataccagtggt 4985  
 4986 gttgaagaaagcaaggataatttggaaacaagcttctgcagatgacaggctaaattttggtgaccgaattttggtaaaagc 5065  
 5066 cccaggttatccatgggtggccgacctgctactgagacgaaaagaactaaaggatagtttgaataactaatagctcattta 5145  
 5146 atgtcttataaagggttttttccctgacttcaattttgcattgggtgaaagaaatagtttaagccattattggat 5225  
 5226 tccgaaatagccaaatttcttgggttcctcaaaagcggaagtctaaagaacttatggaagcttatgaggcttcaaaaactcc 5305  
 5306 tccgtgatttaaggagggaatcttcaccgatgaggaatggaatgattatcagctgtgaggagaagcctaatttttgc 5385  
 5386 aaaaaagaaaatatcatgtggagacatctcttgatgaatcagatgcggagagtatctccagcggatcccttgatgtcaata 5465  
 5466 acttctatttctgaaatgtatgggtcctactgtcgttctgacttctcgtagctctacgcagttaagtgaacaaagggtacc 5544

FIG. 46  
(CONTINUED)

+

1  
 met ser val tyr val val glu leu leu  
 ATG AGT GTG TAC GTC GTC GAG CTG CTC  
 GCCAAGTTCCTGCACTGGCTG  
 10 20  
 arg ser phe phe tyr val thr glu thr thr phe gln lys asn arg  
 AGG TCT TTC TTT TAT GTC ACG GAG ACC ACG TTT CAA AAG AAC AGG  
 30  
 leu phe phe tyr arg lys ser val trp ser lys leu gln ser ile  
 CTC TTT TTC TAC CGG AAG AGT GTC TGG AGC AAG TTG CAA AGC ATT  
 40 50  
 gly ile arg gln his leu lys arg val gln leu arg glu leu ser  
 GGA ATC AGA CAG CAC TTG AAG AGG GTG CAG CTG CGG GAG CTG TCG  
 60  
 glu ala glu val arg gln his arg glu ala arg pro ala leu leu  
 GAA GCA GAG GTC AGG CAG CAT CGG GAA GCC AGG CCC GCC CTG CTG  
 70 80  
 thr ser arg leu arg phe ile pro lys pro asp gly leu arg pro  
 ACG TCC AGA CTC CGC TTC ATC CCC AAG CCT GAC GGG CTG CGG CCG  
 90  
 ile val asn met asp tyr val val gly ala arg thr phe arg arg  
 ATT GTG AAC ATG GAC TAC GTC GTG GGA GCC AGA ACG TTC CGC AGA  
 100 110  
 glu lys ala glu arg leu thr ser arg val lys ala leu phe  
 GAA AAG ARG GCC GAG CGT CTC ACC TCG AGG GTG AAG GCA CTG TTC  
 120  
 ser val leu asn tyr glu arg ala arg arg pro gly leu leu gly  
 AGC GTG CTC AAC TAC GAG CGG GCG CGG CGC CCC GGC CTC CTG GGC  
 130 140  
 ala ser val leu gly leu asp asp ile his arg ala trp arg thr  
 GCC TCT GTG CTG GGC CTG GAC GAT ATC CAC AGG GCC TGG CGC ACC  
 150  
 phe val leu arg val arg ala gln asp pro pro pro glu leu tyr  
 TTC GTG CTG CGT GTG CGG GCC CAG GAC CCG CCG CCT GAG CTG TAC  
 160 170  
 phe val lys val asp val thr gly ala tyr asp thr ile pro gln  
 TTT GTC AAG GTG GAT GTG ACG GGC GCG TAC GAC ACC ATC CCC CAG  
 180  
 asp arg leu thr glu val ile ala ser ile ile lys pro gln asn  
 GAC AGG CTC ACG GAG GTC ATC GCC AGC ATC ATC AAA CCC CAG AAC  
 190 200  
 thr tyr cys val arg arg tyr ala val val gln lys ala ala met  
 ACG TAC TGC GTG CGT CGG TAT GCC GTG GTC CAG AAG GCC GCC ATG

FIG. 47



210  
 gly thr ser ala arg pro ser arg ala thr ser tyr val gln cys  
 GGC ACG TCC GCA AGG CCT TCA AGA GCC ACG TCC TAC GTC CAG TGC  
 220 230  
 gln gly ile pro gln gly ser ile leu ser thr leu leu cys ser  
 CAG GGG ATC CCG CAG GGC TCC ATC CTC TCC ACG CTG CTC TGC AGC  
 240  
 leu cys tyr gly asp met glu asn lys leu phe ala gly ile arg  
 CTG TGC TAC GGC GAC ATG GAG AAC AAG CTG TTT GCG GGG ATT CGG  
 250 260  
 arg asp gly leu leu leu arg leu val asp asp phe leu leu val  
 CGG GAC GGG CTG CTC CTG CGT TTG GTG GAT GAT TTC TTG TTG GTG  
 270  
 thr pro his leu thr his ala lys thr phe leu arg thr leu val  
 ACA CCT CAC CTC ACC CAC GCG AAA ACC TTC CTC AGG ACC CTG GTC  
 280 290  
 arg gly val pro glu tyr gly cys val val asn leu arg lys thr  
 CGA GGT GTC CCT GAG TAT GGC TGC GTG GTG AAC TTG CGG AAG ACA  
 300  
 val val asn phe pro val glu asp glu ala leu gly gly thr ala  
 GTG GTG AAC TTC CCT GTA GAA GAC GAG GCC CTG GGT GGC ACG GCT  
 310 320  
 phe val gln met pro ala his gly leu phe pro trp cys gly leu  
 TTT GTT CAG ATG CCG GCC CAC GGC CTA TTC CCC TGG TGC GGC CTG  
 330  
 leu leu asp thr arg thr leu glu val gln ser asp tyr ser ser  
 CTG CTG GAT ACC CGG ACC CTG GAG GTG CAG AGC GAC TAC TCC AGC  
 340 350  
 tyr ala arg thr ser ile arg ala ser leu thr phe asn arg gly  
 TAT GCC CGG ACC TCC ATC AGA GCC AGT CTC ACC TTC AAC CGC GGC  
 360  
 phe lys ala gly arg asn met arg arg lys leu phe gly val leu  
 TTC AAG GCT GGG AGG AAC ATG CGT CGC AAA CTC TTT GGG GTC TTG  
 370 380  
 arg leu lys cys his ser leu phe leu asp leu gln val asn ser  
 CGG CTG AAG TGT CAC AGC CTG TTT CTG GAT TTG CAG GTG AAC AGC  
 390  
 leu gln thr val cys thr asn ile tyr lys ile leu leu leu gln  
 CTC CAG ACG GTG TGC ACC AAC ATC TAC AAG ATC CTC CTG CTG CAG  
 400 410  
 ala tyr arg phe his ala cys val leu gln leu pro phe his gln  
 GCG TAC AGG TTT CAC GCA TGT GTG CTG CAG CTC CCA TTT CAT CAG

**FIG. 47**  
 (CONTINUED)

+

58/59

```

                                420
gln val trp lys asn pro his phe ser cys ala ser ser leu thr
CAA GTT TGG AAG AAC CCA CAT TTT TCC TGC GCG TCA TCT CTG ACA

430                                440
arg leu pro leu leu leu his pro glu ser gln glu arg arg asp
CGG CTC CCT CTG CTA CTC CAT CCT GAA AGC CAA GAA CGC AGG GAT

                                450
val ala gly gly gln gly arg arg arg pro ser ala leu arg gly
GTC GCT GGG GGC CAA GGG CGC CGC CGG CCC TCT GCC CTC CGA GGC

460                                470
arg ala val ala val pro pro ser ile pro ala gln ala asp ser
CGT GCA GTG GCT GTG CCA CCA AGC ATT CCT GCT CAA GCT GAC TCG

                                480
thr pro cys his leu arg ala thr pro gly val thr gln asp ser
ACA CCG TGT CAC CTA CGT GCC ACT CCT GGG GTC ACT CAG GAC AGC

490                                500
pro asp ala ala glu ser glu ala pro gly asp asp ala asp cys
CCA GAC GCA GCT GAG TCG GAA GCT CCC GGG GAC GAC GCT GAC TGC

                                510
pro gly gly arg ser gln pro gly thr ala leu arg leu gln asp
CCT GGA GGC CGC AGC CAA CCC GGC ACT GCC CTC AGA CTT CAA GAC

520                                530
his pro gly leu met ala thr arg pro gln pro gly arg glu gln
CAT CCT GGA CTG ATG GCC ACC CGC CCA CAG CCA GGC CGA GAG CAG

                                540
thr pro ala ala leu ser arg arg ala tyr thr ser gln gly gly
ACA CCA GCA GCC CTG TCA CGC CGG GCT TAT ACG TCC CAG GGA GGG

550                                560
arg gly gly pro his pro gly leu his arg trp glu ser glu ala
AGG GGC GGC CCA CAC CCA GGC CTG CAC CGC TGG GAG TCT GAG GCC

564
OP
TGA GTGAGTGTTTGGCCGAGGCCTGCATGTCCGGCTGAAGGCTGAGTGTCGGCTGAGGC
CTGAGCGAGTGTCAGCCAAGGGCTGAGTGTCAGCACACCTGCGTTTTCACTTCCCCAC
AGGCTGGCGTTCGGTCCACCCAGGGCCAGCTTTTCCTCACCAGGAGCCCGGCTTCCACT
CCCCACATAGGAATAGTCCATCCCCAGATTGCGCCATTGTTTACCCTTCGCCCTGCCTTCC
TTTGCCTTCCACCCCAACCATTCAGGTGGAGACCCTGAGAAGGACCCTGGGAGCTTTGGG
AATTTGGAGTGACCAAAGGTGTGCCCTGTACACAGGCGAGGACCCTGCACCTGGATGGGG
GTCCCTGTGGGTCAAATTTGGGGGGAGGTGCTGTGGGAGTAAAATACTGAATATATGAGTT
TTTCAGTTTTTGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

```

**FIG. 47**  
(CONTINUED)

+

+

59/59

Motif -1  
 Ep p123 ...LVVSLIRCFFYVTEQQKSYSKT...  
 Sp Tez1 ...FIIPILQSFFYITESDLNRNT...  
 Sc Est2 ...LIPKIIQTFFYCTEISSTVTIV...  
 Hs TCP1 ...YVVELLRSFFYVTETTTFQKNRL...  
 consensus FFY TE

Motif 0  
 Ep p123 ...KSLGFAPGKLRRLIPKKT--TFRPIMTFNKKIV...  
 Sp Tez1 ...QKTTLPPAVIRLLPKKN--TFRLITNLRKRFL...  
 Sc Est2 ...TLSNFNHSMRIIPKKSNEFRIIAIPCRGAD...  
 Hs TCP1 ...ARPALLTSRLRFIPKPD--GLRPIVNMDYVVG...  
 consensus R PK R I

Motif A  
 Ep p123 ...PKLFFATMDIEKCYDSVNREKLSTFLK...  
 Sp Tez1 ...RKKYFVRIDIKSCYDRIKQDLMFRIVK...  
 Sc Est2 ...PELYFMKFDVKSCYDSIPRMECMRILK...  
 Hs TCP1 ...PELYFVKVDVTGAYDTIPQDRLTEVIA...//...  
 consensus F D YD

Motif B  
 Ep p123 ...NGKFYKQTKGIPQGLCVSSILSSFYYA...  
 Sp Tez1 ...GNSQYLQKVGIPQGSILSSFLCHFYME...  
 Sc Est2 ...EDKCYIREDGLFQGSSLSAPIVDLVYD...  
 Hs TCP1 ...RATSYVQCQGIPQGSILSTLLCSLCYG...  
 consensus G QG S

Motif C  
 Ep p123 ...PNVNLLMRLTDDYLLITTQENN...  
 Sp Tez1 ...KKGSVLLRVVDDFLFITVNKKD...  
 Sc Est2 ...SQDTLILKLADDFLIISTDQQQ...  
 Hs TCP1 ...RRDGLLLRLVDDFLVTPHLTH...  
 consensus DD L

Motif D  
 Ep p123 ...NVSRENGFKFNMKKL...  
 Sp Tez1 ...LNLSLRGFKEKHNFS...  
 Sc Est2 ...KKLAMGGFQKYNAKA...  
 Hs TCP1 ...LRTLVRGVPEYGCVV...  
 consensus G

FIG. 48

+